Speech discrimination after early exposure to pulsed-noise or speech

Kamalini G. Ranasinghe[1], **Ryan S. Carraway**[1], **Michael S. Borland**[1], **Nicole A. Moreno**[1], **Elizabeth A. Hanacik**[1], **Robert S. Miller**[1], and **Michael P Kilgard**[1]

[1]School of Behavioral and Brain Sciences, GR41 The University of Texas at Dallas, 800 West Campbell Road, Richardson, Texas 75080-3021

Abstract

Early experience of structured inputs and complex sound features generate lasting changes in tonotopy and receptive field properties of primary auditory cortex (A1). In this study we tested whether these changes are severe enough to alter neural representations and behavioral discrimination of speech. We exposed two groups of rat pups during the critical period of auditory development to pulsed noise or speech. Both groups of rats were trained to discriminate speech sounds when they were young adults, and anesthetized neural responses were recorded from A1. The representation of speech in A1 and behavioral discrimination of speech remained robust to altered spectral and temporal characteristics of A1 neurons after pulsed-noise exposure. Exposure to passive speech during early development provided no added advantage in speech sound processing. Speech training increased A1 neuronal firing rate for speech stimuli in naïve rats, but did not increase responses in rats that experienced early exposure to pulsed noise or speech. Our results suggest that speech sound processing is resistant to changes in simple neural response properties caused by manipulating early acoustic environment.

Keywords

pulsed-noise; early acoustic experience; primary auditory cortex; speech discrimination; plasticity; critical period

1. INTRODUCTION

Early auditory experience alters the development of primary auditory cortex (A1). Exposure to structured inputs during critical period of auditory development induces changes of receptive field properties and temporal response dynamics that endure into adulthood (de Villers-Sidani et al., 2007; Han et al., 2007; Insanally et al., 2010; Insanally et al., 2009; Kim et al., 2009; Zhang et al., 2001; Zhang et al., 2002; Zhou et al., 2007; Zhou et al., 2008; Zhou et al., 2009). Exposure to moderate intensity pulsed noise during development results in deteriorated tonotopicity, broader tone frequency tuning and degraded cortical temporal processing manifested by weaker responses to rapid tone trains (Zhang et al., 2002; Zhou et al., 2008; Zhou et al., 2009). Exposure to pulsed tones during development broadens tone

frequency tuning and results in an expansion of the A1 representations of the exposed tone frequency (de Villers-Sidani et al., 2007; Zhang et al., 2001). Exposure to downward FM sweeps during development gives rise to greater selectivity for the sweep rate and direction of the exposed stimulus (Insanally et al., 2009). It was suggested that cortical plasticity induced by manipulating early acoustic inputs could impair the processing of complex sounds such as speech (Chang et al., 2005; Zhou et al., 2007; Zhou et al., 2008; Zhou et al., 2009). Given the profound effects of native language exposure on speech perception in later life, the early acoustic experience of speech itself is likely to play an important role in shaping the auditory cortex (Kuhl, 2004; Kuhl et al., 2003; Kuhl et al., 1997a; Kuhl et al., 1997b). Neural circuits shaped by unique language experience during infancy have been suggested as a primary determinant factor in behavioral ability to discriminate speech (Cheour et al., 1998; Kuhl, 2010; Naatanen et al., 1997). However, the degree of developmental auditory cortex plasticity that is necessary to produce a significant effect on neural representation and behavioral discrimination of speech has not been tested.

Earlier studies have shown that neural activity patterns generated in rat A1 are significantly correlated with behavioral discrimination of speech (Engineer et al., 2008). Rats discriminated easily between consonant pairs that generated distinct spatiotemporal patterns and did poorly between pairs that generated similar patterns. Recent evidence from intraoperative recordings in humans have shown that representation of speech signals in posterior superior temporal gyri significantly predict the psychophysical behavior of speech recognition by humans (Chang et al., 2010). These observations suggest that A1 representations play a key role in speech sound processing and subsequent behavioral outcome of speech perception. In this study we tested the hypothesis that, changes of spectral and temporal selectivity of A1 neurons induced by early acoustic environmental patterns are severe enough to influence the behavioral and neural discrimination of speech. We exposed rat pups to two different manipulations of early auditory experience (i.e. exposure to pulsed-noise or speech) and later trained them to discriminate speech sounds. We examined the anesthetized A1 neural activity patterns of pulsed-noise-reared and speech-reared rats before and after speech training.

2. METHODS

2.1 Early exposure

We exposed two groups of Sprague-Dawley rat pups during their postnatal development from day 9 (P9) through day 38 (P38) to either a pulsed noise stimulus or to speech sounds (Figure 1). The pulsed-noise-rearing consisted of exposure to fifty- millisecond noise pulses (5-ms ramps) at 6 pulses per second pps with 1-s intervals in between. The speech-rearing exposure consisted of 13 CVC (consonant-vowel-consonant) speech syllables, each 1.5 seconds long. To match the most ethological rates of presentation, the 13 speech sounds were randomly interleaved with 13 silent wave files of the same duration and a single wave (speech or silent) was presented once every 2 seconds in random order. These 13 syllables consisted of the same CVC syllables that rats were trained to discriminate in later life (see methods below). Both pulsed-noise and speech sounds contained energy levels below 25 kHz. All sounds were played through an iPOD connected to a speaker placed ~15 cm above the rats and with 65-dB sound pressure level. During the exposure period litters of rat pups and their mothers were placed in a sound shielded chamber. A reverse 12-hour light/ 12-hour dark cycle was maintained to provide the normal physiological balance during awake and sleep states. Nearly half of the control group of rats was reared in an identical sound shielded chamber that did not play any sounds from P9 to P38. The rest of the control rats were reared (from P9 to P38) in standard housing room environments. The tuning curve properties of the rats reared in the sound shielded chamber without any sounds were not significantly different from the rats reared in the standard housing room condition. For the

analysis, therefore, we combined the two sets of control rats into a single control group. Only the female pups were used in the experiments. All experimental procedures were approved by the Animal Care and Use Committee at the University of Texas at Dallas.

2.2 Speech training and analysis

We trained 50 days (P50) old juvenile rats to discriminate 6 consonant pairs and 6 vowel pairs. The 6 consonant pairs included Dad/Bad, Dad/Gad, Dad/Tad, Sad/Shad, Sad/Had, and Sad/Dad. The 6 vowel pairs included Dad/Dud, Dad/Did, Dad/Dayed, Sad/Seed, Sad/Sead, and Sad/Sood. The spectrogram illustrations of these stimuli have been published elsewhere (Engineer et al., 2008; Perez et al., 2012, in press). 12 rats from each group of pulsed-noise-reared, speech-reared and control group were trained to discriminate speech in an operant training paradigm. Rats pressed a lever for a food reward for a target stimulus (CS plus) and rejected the lever for non-target stimuli (CS minus). Half the rats in each group (6 rats) were trained on the target 'Dad' (non-targets: Bad, Gad, Tad, Dud, Did, Dayed) and the other half were trained on the target 'Sad' (non-targets: Shad, Had, Dad, Seed, Sead, Sood). Thus each rat performed three consonant discrimination tasks and three vowel discrimination tasks. The speech sounds were spoken by a native female English speaker inside a double walled sound-proof booth. To better match the rat hearing range, the fundamental frequency of the spectral envelope of each recorded syllable was shifted up in frequency by a factor of two using the STRAIGHT vocoder (Kawahara, 1997). The intensity of each speech sound was adjusted so that the intensity of the loudest 100ms of the syllable was at 60 dB sound pressure level (SPL).

The operant training chamber consisted of a double walled, sound shielded booth containing a cage with a lever, a pellet receptacle and a house light. A pellet dispenser mounted outside the booth was connected to the pellet receptacle via a tube. A calibrated speaker was mounted approximately 20 cm from the most common location of the rat's head position while inside the booth (the mid point between the lever and the pellet receptacle). Rats were food deprived during the training days to motivate behavior, but were fed otherwise to maintain between 85%-100% of their ideal body weight. Each rat was placed in the operant training booth for one-hour session, twice daily for five days per week.

The training began with a brief shaping period to teach the rat to press the lever in order to receive a food pellet reward. Each time the rat was in close proximity to the lever, it heard the target sound, and received a sugar pellet reward. Eventually the rat learned to press the lever without assistance. The shaping continued until the rat was able to earn a minimum of 120 pellets per session, which lasted on average 3-5 days. During the next stage of training the rat began a detection task where it was trained to press the lever when the target sound was presented. Initially the sound was presented every 10 seconds and the rat received a sugar pellet reward for pressing the lever within 8 seconds. The sound interval was gradually reduced to 6 seconds and the lever press window to 3 seconds. The target sound was randomly interleaved with silent catch trials during each training session to prevent rats pressing the lever habitually. Performance of the rat was monitored using the d-prime (d') value, which is a measure of discriminability according to signal detection theory (Green et al., 1989). The detection training continued until the rats reached a d' 1.5 for six sessions. This portion of training lasted for ~ 15 days (~ 30 sessions). After completing the detection phase, the rats learned to discriminate the target sound from the non-targets. In a given one hour session each rat discriminated the target sound against the six non-target sounds that were presented randomly interleaved with silent catch trials. Rats were rewarded with a sugar pellet when they pressed the lever within 3 seconds following the target stimulus. Pressing the lever for a non-target stimulus resulted in a 6 second time-out in which the training program paused and the house light was turned off. The discrimination training consisted of 20 sessions (10 days). The final percent correct discrimination performance for

each task was computed using the number of hits/misses/false-alarm and correct-rejection rates recorded during the last four sessions for each rat. The rats aged between 80 – 100 days at the completion of the behavioral training. Eight rats from each group of pulsed-noise-reared, speech-reared, and control group (each eight consisting of four from 'Dad' target group and four from 'Sad' target group) were then used to examine the electrophysiology responses of A1 neurons. The physiology recordings were done within a day following the completion of the behavior training.

2.3 Electrophysiology and analysis

Multiunit clusters from right A1 were recorded under anesthesia from behavior trained rats and untrained rats of each group (pulsed-noise-reared, speech-reared, control). We recorded 242 A1 sites of pulsed-noise-reared untrained (n=6), 335 A1 sites of pulsed-noise-reared trained (n=8), 155 A1 sites of speech-reared untrained (n=4), 320 A1 sites of speech-reared trained (n=8), 281 A1 sites of control untrained (n=9), and 290 A1 sites of control trained (n=8) rats. Rats were aged from 80-100 days at the time of physiology recordings (Figure 1). Anesthesia was induced with a single dose of intra-peritoneal pentobarbital (50 mg/kg). To maintain anesthesia, rats received supplemental dilute pentobarbital (8 mg/ml) every ½ to 1 hour, subcutaneously. The supplemental dosage was adjusted as needed to maintain areflexia. A 1:1 mixture of dextrose and Ringers lactate was given every ½ hour to 1 hour to prevent dehydration. Heart rate, respiratory rate, Oxygen saturation and body temperature were monitored throughout the experiment. A1 responses were recorded with four Parylene-coated tungsten micro-electrodes (1-2 MΩ, FHC). The electrodes were lowered simultaneously to 600 -700 μm below the surface of the right cortex corresponding to layer IV. Each electrode penetration site was marked on a map using cortical blood vessels as landmarks. At each recording site 25 ms tones were played at 90 logarithmically spaced (1-47 kHz) frequencies at 16 intensities (0-75 dB SPL), to determine the frequency-intensity tuning curves of each site. After recording the tuning curve for each site speech stimuli consisting of 13 speech sounds were presented. Each stimulus was played 20 times per site, in random order and once every 2000 ms, at 65 dB SPL. After the presentation of tones and speech, we played noise-burst-trains of variable repetition rates to determine temporal modulation functions of A1 neurons. Eight different repetition rates (2, 4, 7, 10, 12.5, 15, 17.5, 20 pps) were used and each noise-burst-train was presented 20 times, at 65 dB SPL. The repetition rates were randomly interleaved, and 200 ms of silence separated each train. All stimuli were played from a calibrated speaker mounted approximately 10 cm from the base of the left pinna. Stimulus generation, data acquisition and spike sorting was performed with Tucker-Davis hardware (RP2.1 and RX5) and software (Brainware).

We quantified the receptive field properties for each group, including threshold, bandwidths, latency, spontaneous firing, and response strength to tones. Threshold was defined as the lowest intensity that evoked a response at the characteristic frequency for each site. Bandwidths were measured in octaves as the frequency range that evoked a response to 10, 20, 30, and 40 dB above the threshold. We used customized MATLAB software to analyze these properties and the analysis was done by a blind observer. The percent of A1 responding to tones presented at different intensity levels was calculated for pulsed-noise-reared, speech-reared, control rats. The percent of cortex was compared between different exposure groups using measures described in previous studies (Kilgard et al., 1998a; Puckett et al., 2007). In this analysis Voronoi tessellation was used to transform the discretely sampled cortical surface into a continuous map using the assumption that each point on the map has the response characteristics of the nearest recording site. The percentage of the cortical area responding to each tone was estimated as the sum of the areas of tessellations from A1 sites that included the tone, as a fraction of the total area of A1.

We analyzed the responses to noise burst trains to document the response synchronization and response strength of A1 neurons. Response synchronization was quantified using vector strength and Rayleigh statistic measures. Vector strength quantifies the degree of synchronization between action potentials and repeated sounds. Vector strengths equal to a value of one demonstrates perfect synchronization, whereas no synchronization would result in a value of zero. We also computed the Rayleigh statistic, which is a circular statistic to assess the significance of vector strength values. Rayleigh statistic values greater than 13.8 indicate statistically significant ($P < 0.01$) phase locking. Response strength of A1 firing to noise burst trains was quantified by temporal modulation transfer function. We analyzed the temporal modulation transfer function for the average number of responses (above spontaneous firing) following the second through sixth bursts of each train as a function of the repetition rate. The methods used in this analysis were similar to those described in previous studies (Kilgard et al., 1998b; Pandya et al., 2008).

We created neurograms for each rat, using the A1 evoked activity of the onset consonant (first 40 ms of the CVC syllable) for the seven consonants used in the study (D, B, G, T, S, SH, H). Each neurogram consisted of responses from A1 recording sites (per rat) on the y axis ordered from low to high characteristic frequency, and time (0 – 40 ms) on x axis providing 1 ms precise spike timing information. Neurogram response at each site consisted of the average of the 20 repetitions of a given stimulus played at each site. Neural dissimilarity between different consonant pairs was computed using the Euclidean distance between neurograms. The Euclidean distance between a given neurogram pair is the square root of the mean of the squared differences between the firing rates for each bin averaged across the number of sites, for a given rat. To maintain an adequate representation of low, mid and high frequency tuned neurons for each rat, we only used data from rats that had more than 25 A1sites for this analysis (n=6, n=4, n=8 for pulsed-noise-reared, speech-reared and control rats respectively). The Euclidean distance per group (pulsed-noise-reared, speech-reared, control) was computed by averaging across the rats in each group. The measures used in analyzing neural dissimilarity between consonant pairs were similar to those described in previous experiments (Engineer et al., 2008)

In contrast to rapid spectral changes of consonants, vowels represent relatively stable spectral patterns. Previous neurophysiological studies have reported that different vowel sounds generate distinct spatial activity patterns in the central auditory system (Delgutte et al., 1984; Ohl et al., 1997; Sachs et al., 1979) These observations suggest that the number of spikes (spike count) generated by A1 neurons as a function of characteristic frequency could be used to discriminate vowel sounds. We created spike-count profiles for each rat using the A1 evoked activity for the first 350 ms after voicing of the initial consonant of CVC syllables depicting the neural activity pattern for the vowel. A spike-count profile for a given stimulus consisted of the total number of spikes fired plotted on y axis, by each recording site arranged from low to high frequency on x axis. The neural response at each site consisted of the average of the 20 repetitions of a given stimulus played at that site. The Euclidean distance between spike-count profiles were computed as a measure of neural dissimilarity between vowels. The Euclidean distance between a given spike-count pair is the square root of the mean of the squared differences between the firing rates during the 350ms window (350 ms bin) averaged across the number of sites for a given rat. Similar to consonant analysis, to maintain an adequate representation of low, mid and high frequency tuned neurons for each rat we only used data from rats that had more than 25 A1sites for this analysis. The Euclidean distance per group (pulsed-noise-reared, speech-reared, control) was computed by averaging across the rats in each group. The analysis criteria of the neural evoked activity of spike count for vowels were similar to the measures described in previous studies (Perez et al., 2012, in press).

A neural classifier (Engineer et al., 2008; Foffani et al., 2004) was used to calculate the neural discrimination between speech stimuli in units of percent correct by using the individual sweeps. The neural classifier randomly picked a single sweep (out of 20) and made a template from 19 sweeps (the sweep being considered was excluded) for each speech sound. The sweep under consideration was then compared to the templates using the Euclidean distance as a measure of neural similarity. The classifier determined that the single sweep response under consideration was generated by the sound whose template was closest to it in Euclidean space. This model assumed that the average templates are analogous to the neural memory of different sounds. For consonant analysis the classifier used the first 40 ms of the speech evoked responses and considered the precise spike timing information (1 ms temporal resolution). For vowel analysis the classifier used the first 350 ms of the voiced speech evoked activity and considered the total number of spikes fired. The classifier analysis was done for each rat determining the neural discrimination by a single subject. Neural discrimination performance for each group was calculated by averaging across the rats in each group.

3. RESULTS

We documented the changes in temporal and spectral properties of A1 neurons after exposing rat pups to either pulsed-noise or speech sounds from postnatal day 9 to day 38 (P9 - P38). The behavioral and neural discrimination of speech (6 consonant discrimination tasks and 6 vowel discrimination tasks) was examined in exposed juvenile rats (P50 – P100), and compared to those of unexposed control rats. In the subsequent sections we first present the tuning curve properties and speech evoked neural responses of A1 neurons recorded from untrained rats and correlate the neural differences to behavioral discrimination (sections 3.1 through 3.3). Next we present the speech evoked neural responses recorded from both speech trained and untrained rats (section 3.4).

3.1 Early auditory experience modify A1 response characteristics

Previous experiments have shown that early auditory experience of pulsed-noise results in decreased phase locking to tone trains when presented at high repetition rates and increased phase locking at low repetition rates (Zhou et al., 2008; Zhou et al., 2009). It is reasonable to expect that altered synchronization of A1 neurons would alter neural responses to speech, which is a temporally complex signal. We tested the hypothesis that the ability of A1 responses to phase lock to a modulated broadband noise signal will be affected by pulsed-noise or speech exposure. The modulation rates tested span the range of temporal envelope modulations in speech (2 -20 Hz). Vector strength was used to characterize the strength of phase locking to noise burst trains. The pulsed-noise reared rats and speech-reared rats showed significantly larger vector strengths at higher repetition rates (i.e. 7 - 20 Hz), compared to control rats. For example at 20 Hz, speech-reared rats showed a 43%(ANOVA with post hoc Dunnett's P<0.05) larger vector strength and the pulsed-noise-reared rats showed a 16% (ANOVA with post hoc Dunnett's P<0.05) larger vector strength (Figure 2a) compared to controls. At lower repetition rates (i.e. 2 - 4 Hz), the pulsed-noise-reared rats showed significantly smaller vector strengths, compared to control rats. For example at 2 Hz, pulsed-noise-reared rats showed a 28%(ANOVA with post hoc Dunnett's P<0.05) smaller vector strength compared to control rats (Figure 2a). Rayleigh statistic values indicated significantly phase locked neural activity across all repetition rates (P<0.01) (Figure 2b). Collectively, these results support the hypothesis that early auditory experience significantly changes the response synchronization of A1 neurons to temporal modulations of a broadband stimulus. However, in contrast to the decreased phase locking ability at higher rates and increased phase locking ability at lower rates of tone trains as described in previous studies (Zhou et al., 2008; Zhou et al., 2009), we observed an increased phase locking ability at higher rates and decreased phase clocking ability at lower rates of noise

bursts. This difference in temporal plasticity following pulsed-noise rearing is likely to be due to the difference in the sound used to quantify temporal responses (i.e. noise burst trains versus tone trains). Our results show that early auditory exposure to structured complex signals (i.e. pulsed-noise or speech) significantly increases the temporal fidelity of A1 firing at higher repetition rates of broadband stimuli. We predict that altered ability of A1 phase locking will be severe enough to modify the speech sound processing in pulsed-noise-reared and speech-reared rats.

Previous studies showed that response strength of A1 neurons to temporally modulated stimuli are significantly influenced by early auditory experience. Fast tone trains (i.e. above 7 Hz) generate fewer action potentials in pulsed-noise-reared rats' than in controls, and slow tone trains (i.e. 2 – 4 Hz) generate more action potentials than in controls (Zhou et al., 2008; Zhou et al., 2009). We predicted that the response strength to a temporally modulated broadband stimulus would be altered by auditory experience of pulsed-noise or speech during development. The number of spikes evoked by a train of 6 noise bursts presented at different rates was documented in pulsed-noise-reared, speech-reared and control rats. The control rats evoked the same or greater number of spikes per each noise burst in the train compared to the first burst, thus maintaining a normalized spike rate of 1 up to repetition rates of 7 Hz (Figure 2c). In contrast, pulsed-noise reared rats' and speech-reared rats' A1 neurons maintained a normalized firing rate of 1 up to repetition rate of 10 Hz (normalized spike rate=1.18± 0.05 and 1.14±0.08, for pulsed-noise and speech respectively; ANOVA and post hoc Dunnett's, P<0.05)(Figure 2c). At slower rates of presentation (i.e. 2 – 4 Hz) the pulsed-noise reared rats evoked significantly fewer spikes per noise burst, compared to control rats. For example at 4 Hz, the pulsed-noise-reared rats maintained a significantly low normalized firing response of 1.6±0.05, as opposed to 2.1±0.12, of control rats (ANOVA and post hoc Dunnett's, P<0.05, Figure 2c). These findings support the prediction that early experience of pulsed-noise and speech have significant effects on A1 response strength to a temporally modulated broadband signal. The increased temporal following rate for noise burst trains presented at 10 Hz after pulsed-noise exposure in the current study differs from the reduced temporal following rate for tone trains in pulsed-noise-reared rats reported in earlier studies(Zhou et al., 2008; Zhou et al., 2009). It is likely that this difference in temporal following rate is due to the difference in the stimulus (i.e. noise burst trains versus tone trains) used to quantify temporal responses. We found that early exposure to structured inputs (i.e. pulsed-noise or speech) makes A1 neurons better at following temporally modulated broadband stimuli at 10 Hz. Given that speech is primarily modulated at rates between 2 – 20Hz, we predicted that A1 responses to speech stimuli would be altered in exposed rats.

Earlier reports on rearing rat pups in different environmental conditions have shown that continuous noise delays the emergence of tonotopic functional organization of A1, whereas modulated noise results in degraded tonotopicity of A1 (Chang et al., 2005; de Villers-Sidani et al., 2007; Zhang et al., 2001; Zhang et al., 2002). We examined the frequency representation of A1 neurons after exposing rat pups to pulsed-noise or speech sounds during early development. We documented the percentage of neurons responding to a given frequency-intensity combination in each animal and estimated the means for each exposure group. Pulsed-noise-reared rats showed a significantly increased proportion of neurons responding to low frequencies (i.e.1-3 kHz) compared to control rats, and a significantly decreased proportion of neurons responding to high frequencies (i.e.16 – 32kHz). For example, 62±8% of pulsed-noise reared rats' A1 neurons responded to a 2 kHz tone played at 45 dB, as opposed to 46±6% of A1 neurons in control rats (Figure 3a and b, t(13)=2.2, P<0.05), but only 38±4% of pulsed-noise reared rats' A1 neurons responded to a 16 kHz tone played at 45 dB, as opposed to 48±6% of A1 neurons in control rats (Figure 3a,b; t(13)=2.4, P<0.05). The enhanced low frequency and reduced high frequency tuning of A1

neurons in pulsed-noise-reared rats is consistent with previous reports (Zhang et al., 2002). Speech rearing, on the other hand, did not reliably alter the A1 response to tones (Figure 3a and c). This result shows that early experience of speech sounds does not produce significant tonotopic changes, despite containing complex temporally structured inputs.

Exposure to pulsed-noise as well as frequency modulated sweeps during early development has been shown to change the receptive field size of A1 neurons (Insanally et al., 2010; Insanally et al., 2009; Zhang et al., 2002; Zhou et al., 2009). When reared in pulsed-noise or downward FM sweeps tuning bandwidths of A1 neurons were either broadened or narrowed depending on the timing of exposure. Exposure during early postnatal period (i.e. P16 – P23) broadened the tuning bandwidths where as exposure during late postnatal period (i.e. P32 – P39) narrowed the bandwidths (Insanally et al., 2010; Insanally et al., 2009). We compared the tuning bandwidths of A1 neurons of control rats to those of rat pups exposed to pulsed-noise or speech from P9 to P38. Pulsed-noise rearing resulted in a modest broadening of bandwidths in A1 neurons. For example, the tuning bandwidths showed a 13% broadening at 20dB above threshold (BW20) in pulsed-noise rats compared to control rats (Figure 3d, $F_{(2,15)}=5.33$, $P<0.05$; pulsed-noise-reared=1.7 ± 0.05 oct, control=1.5± 0.04 oct, 95% confidence interval 0.00977 0.36813, $P< 0.05$, Dunnett's). As opposed to modest broadening of bandwidths by pulsed-noise, speech rearing did not change the tuning bandwidths of A1 neurons (Figure 3d, $F_{(2,15)}=5.33$, $P<0.05$; speech-reared=1.5 ± 0.1 oct, control=1.5± 0.04 oct, 95% confidence interval -0.23119 0.17514, P=NS, Dunnett's). Consistent with previous reports (Insanally et al., 2010; Insanally et al., 2009; Zhang et al., 2002; Zhou et al., 2008) pulsed-noise rearing and speech rearing did not change the onset latencies (ANOVA, $F_{(2,15)}=2.18$, $P=0.15$) or sound intensity thresholds (ANOVA, $F_{(2,15)}=0.6$, $P=0.56$) of A1 neurons (Figure 3e,f). Collectively, we observed that pulsed-noise exposure during P9 to P38 generate a modest increase in receptive field size of A1 neurons while speech exposure during the same time window does not change receptive field size.

In summary, early auditory experience of pulsed-noise and speech increased A1 response synchronization at faster rates of a modulated broadband signal and increased response strengths when presented at 10 Hz. Additionally, early experience of pulsed-noise produced a significant increase in the proportion of neurons responding to low frequency tones plus a modest broadening of receptive field size of A1 neurons. At this point we hypothesized two possible outcomes for speech processing in rats reared in pulsed-noise or speech. First, speech processing may be highly sensitive to altered temporal and spectral properties of A1 neurons. Evidence in support for this hypothesis includes the profound effects on speech and language proficiency in children with significant hearing impairment during infancy (Nicholas et al., 2006; Svirsky et al., 2004). The second hypothesis is that speech processing would remain robust to altered spectral and temporal characteristics. Evidence supporting this includes the robust perception of spectrally and temporally degraded speech (i.e. listening through a cochlear implant) or speech in noise, suggesting that sensory patterns generated by degraded acoustic features can provide highly accurate discrimination of speech (Fishman et al., 1997; Miller et al., 1955; Shannon et al., 1995).

3.2 Pulsed-noise or speech exposure does not change behavioral speech discrimination ability

We compared the behavioral discrimination ability of six consonant discrimination tasks and six vowel discrimination tasks in pulsed-noise-reared and speech-reared rats to that of control rats. The pulsed-noise-reared rats and speech-reared rats learned the discrimination task as rapidly as control rats. For example on their 3rd day of training all three groups (n=12 in each) reached an average d-prime (d') of 1 or above (Figure 4a, pulsed-noise-reared d'=1 ± 0.3; speech-reared d'= 1±0.1; control d'=1 ± 0.1). The average performance accuracy

achieved on the last day of training was also not different in pulsed-noise reared rats (d'=1.6±0.1) and speech-reared rats (d'=1.5±0.1) from that of control rats (1.5±0.1; F(2,33)=0.07, MSE=0.16, P=0.93)(Figure 4a). These findings support the hypothesis that behavioral discrimination of speech is resistant to altered temporal and spectral properties of auditory neurons.

Having demonstrated that overall behavioral discrimination ability of speech remains unchanged after altered early auditory experience, we sought to evaluate whether identification of specific acoustic properties have been affected in pulsed-noise-reared and speech-reared rats. Compared to steady state acoustic properties of vowels, consonants consist of dynamic acoustic features and sharper category boundaries (Delgutte, 1984; Delgutte et al., 1984; Pisoni, 1973). Given that pulsed-noise-rearing and speech rearing produced an increased phase locking at fast temporal modulations of broadband stimuli (Figure 2a), a likely expectation would be that neurons might respond better to rapid acoustic transients of consonants than to slow acoustic transients of vowels. There are two possible behavioral effects following this hypothesis. First, pulsed-noise and speech exposure may provide a significant advantage on discriminating consonants than vowels, compared to control rats. Second, the performance accuracy of consonant discrimination may be significantly better in exposed rats compared to control rats. We examined the behavioral performance of six consonant discrimination tasks and six vowel discrimination tasks in control, pulsed-noise-reared and speech-reared rats. All three groups of rats discriminated consonants better than vowels (81±2% vs. 71±2%, 82±2 vs. 72±1 and 80±2 vs. 72±1, in control, pulsed-noise-reared and speech-reared rats P<0.001 for all three groups). There was no significant advantage of consonant discrimination versus vowel discrimination in pulsed-noise-reared and speech-reared rats compared to control rats (ANOVA F(2,33)=0.46, MSE=14.9, P=0.63). The discrimination accuracy achieved by pulsed-noise-reared rats and speech-reared rats at the end of the training period was similar to those achieved by control rats, for both consonants and vowels (ANOVA F(2,33)=0.33, MSE=27.7, P=0.72 and F(2,33)=0.43, MSE=31.9, P=0.66 for vowels and consonants respectively, Figure 4b and 4c). These observations suggest that behavioral ability to discriminate fast as well as slow acoustic transients of speech is not changed by the degree of alteration in temporal dynamics generated by pulsed-noise rearing and speech rearing.

Earlier studies reported that neurons tuned to different characteristic frequencies vary in their contribution to discriminate consonants with distinct articulatory features. For example, both low and high frequency tuned neurons provide information about stop consonants while only high frequency neurons provide information about fricatives (Engineer, 2008). Given that pulsed-noise rearing produces a significant reduction in the proportion neurons responding to higher frequencies I.e. above 16 Hz, Figure 3b), a likely prediction would be an impaired ability to discriminate fricatives in pulsed-noise-reared rats. However, we found that pulsed-noise-reared rats discriminated fricatives in a similar degree of accuracy as in control rats. For example pulsed-noise-reared rats were able to correctly discriminate 'Sad' from 'SHad' on 88±5% of trials, which was not different from 85±3% by control rats. These results indicate that discrimination of fine acoustic characteristics of speech remains resistant to the degree of altered A1 tonotopicity generated by pulsed-noise-rearing.

Collectively, the performance by exposed rats demonstrates that behavioral discrimination of speech remains robust to changes in spectral and temporal response properties of A1 neurons caused by altered sensory experience during early development. We predict that speech evoked neural activity patterns in pulsed-noise and speech-reared rats would contain sufficient neural differences to support discrimination, despite the altered response characteristics (i.e. response synchronization, response strength and tonotopicity). To test this prediction we recorded the A1 neural activity patterns generated by speech sounds and

compared the neural discrimination ability of pulsed-noise-reared and speech-reared to control rats.

3.3 Pulsed-noise or speech exposure does not change neural discrimination of speech

As reported in previous studies (Engineer et al., 2008) each consonant generated a unique pattern of activity. For example, in control rats, /D/ evoked early high frequency neuronal firing followed by low frequency neurons, while /B/ evoked early low frequency neuronal firing followed by high frequency neurons (Figure 5a, top row). The neural responses for consonants generated in pulsed-noise-reared rats were similar to control rats (Figure 5a, middle row). Previous studies reported that the neural differences generated by consonants as measured by the Euclidean distance of precise spike timing activity (i.e 1 ms bins) in A1 neurons are highly correlated with behavioral discrimination ability of consonants by rats (Engineer et al., 2008; Perez et al., 2012, in press). In the current study we estimated the Euclidean distance between neurograms as a measure of neural difference between consonant pairs. We compared the neural differences generated by A1 neurons of untrained rats to the behavioral discrimination patterns observed in speech training. Consistent with previous studies (Engineer et al., 2008), the consonant pairs that rats discriminated with higher accuracy generated more distinct activity and therefore larger Euclidean distances between the neural responses, compared to pairs which rats discriminated poorly. For example /D/ vs /B/ generated more distinctive patterns compared to /S/ vs /SH/ (Figure 5a top row). As a consequence, /D/ vs /B/ produced larger Euclidean distance (202 ± 10) compared to /S/ vs /SH/ (106 ± 7)(Figure 5b top row), and rats discriminated /D/ vs /B/ easily than /D/ vs /S/ (Figure 4b). The neural differences generated by pulsed-noise-reared rats were similar to control rats (Figure 5b top and middle rows). The average Euclidean distance across all consonant pairs (i.e. 21 pairs) in pulsed-noise-reared rats was not significantly different from control rats (Euclidean distance=172 ± 13, 161 ± 8 for pulsed-noise-reared and control rats respectively; ANOVA $F(2,16)=0.71$, $P=0.50$). We then further quantified the neural discrimination ability of a single sweep of stimulus presentation, which allowed us to calculate the neural discrimination in units of percent correct. This analysis was accomplished with the use of PSTH classifier (see methods). The average neural discrimination ability across all consonant pairs (i.e. 21 pairs) by pulsed-noise reared rats ($76\pm1\%$) was not different from that of control rats ($75\pm1\%$)(ANOVA $F(2,16)=0.27$, $P=0.77$). These results show that A1 neurons in pulsed-noise-reared rats are able to encode sufficient neural differences to produce effective discrimination, thus supporting the earlier prediction that speech processing is resistant to modest alterations of temporal and spectral properties of auditory neurons.

The spatiotemporal activity patterns generated by A1 neurons of speech-reared rats showed similar patterns to those of control rats (Figure 5a, top and bottom rows). The Euclidean distance measures between consonant pairs appeared to be in the same range of control rats (Figure 5b, top and bottom rows). The average Euclidean distance between consonant pairs and the ability of the neural classifier to discriminate consonants was not significantly affected by speech rearing. For example, the neural classifier reached an average accuracy of $76\pm2\%$ in consonant discrimination, which was not different from the control rats' neural classifier performance of $75\pm1\%$ (ANOVA $F(2,16)=0.27$, $P=0.77$). These results suggest that early acoustic experience of speech sounds (and the accompanying increase in phase locking) does not provide an added advantage for neural processing of consonants. Further, these results support our hypothesis of robust sensory discrimination of speech by showing that the altered temporal dynamics of A1 neurons after speech exposure do not degrade the neural differences of speech contrasts.

The number of spikes generated by A1 neurons when plotted against their best frequency, produced unique spatial patterns for each syllable containing a different vowel. For example,

in control rats, 'Sad' generated relatively higher firing rates in low and mid frequency-tuned neurons compared to 'Seed' (Figure 6a top left). The vowel pairs that exist far from each other in vowel-space generated more distinct spatial patterns compared to those that lie closer to each other. For example 'Sad' and 'Seed' generated more distinct spatial patterns and thus a larger Euclidean distance compared to 'Sad' and 'Sead' (Figure 6a,b top rows; Euclidean distance=9±0.8 and 8±0.6 for Sad vs Seed and sad vs Sead respectively). We computed the Euclidean distance between the spatial patterns generated by vowel pairs to quantify the neural difference (Figure 6b, top, middle and bottom rows). The average Euclidean distance generated across 12 vowel pairs in pulsed-noise-reared rats (9±1.2), and speech-reared rats (9±1.0) were not significantly different from that of control rats (8±0.5) (ANOVA $F_{(2,16)}=0.19$, $P=0.83$). The neural classifier quantified the neural discrimination of vowels in units of percent correct. The classifier used the spike count generated in single sweeps separately (first 350 ms of voiced stimulus activity). Neural discrimination of vowels by pulsed-noise-reared rats and speech-reared rats achieved the same degree of accuracy as in control rats (ANOVA $F_{(2,16)}=0.98$, $P=0.40$). These findings support our prediction of neural differences remaining resistant to altered spectral and temporal properties of A1 neurons by demonstrating that spatial patterns generated by steady state acoustic characteristics of vowels remain robust in pulsed-noise-reared and speech-reared rats.

It is possible that greater broadening of receptive field size in A1 neurons could impair speech processing ability. Earlier studies with shorter durations of pulsed-noise exposure (Zhang et al., 2002) reported a larger increase of tuning bandwidth compared to the increase observed in the current study. Previous studies have also reported that exposure to pulsed-noise during a later developmental window (i.e. P32 - P39) narrows tuning bandwidths (Insanally et al., 2010). Exposure to FM sweeps for a longer period (P9 – P39) prevents the broadening caused by a shorter period (P16 – P23)(Insanally et al., 2009). Although no study has yet reported that longer noise exposure used in this study causing a significantly less receptive field broadening, it is possible that exposure to noise for a shorter period of time would have resulted in more broadening than was used in our study. To test whether greater broadening of receptive field size would result in an impairment of neural discrimination of speech, we examined the neural classifier performance as a function of bandwidth at 20 dB above threshold (BW20). We found that neural discrimination performance was weakly but positively correlated with BW20 ($r^2=0.03$, $P<0.00001$, Slope=0.01). This correlation, which suggests that broader tuning bandwidth is associated with slightly better neural discrimination ability, indicates that a greater degree of broadening is unlikely to impair speech discrimination.

3.4 Speech training improves A1 responses in control rats but not in pulsed-noise-reared and speech-reared rats

Perceptual learning associated changes have been reported in A1 neurons with regard to different modalities of training including frequency discrimination, sound localization and temporal discrimination tasks (Recanzone et al., 1993; van Wassenhove et al., 2007). We examined the effects of speech training by comparing the A1 speech evoked activity of behavior trained rats to those of untrained counterparts within pulsed-noise-reared, speech-reared and control rats. We measured the number of A1 spikes generated in the first 100 ms of stimulus onset as a measure of speech evoked activity. A two-way ANOVA revealed significant effect of type of early auditory experience ($F_{(2,37)}=4.2$, $P<0.05$) and significant interaction between the type of early auditory experience and speech training ($F_{(2,37)}=3.6$, $P<0.05$) on A1 speech evoked activity. Speech training induced a significant increase in the response strength of the speech-evoked activity in control rats. For example the number of spikes generated in the first 100 ms of stimulus onset in the trained rats' A1 was significantly increased compared to that of untrained counterparts (Figure 7)(ANOVA with

post hoc Fisher's LSD, trained=47±4, untrained=36±4, P<0.05). Since there was no significant difference in spontaneous activity between trained and untrained groups (trained=11.8±8, untrained=11.5±7; t(15)=8.01 P=0.9), any increase in response strength would increase the signal to noise ratio and therefore neural discrimination. As a consequence of increased response strength therefore the average Euclidean distance measure across the consonant pairs showed a significant increase after training compared to pre-training in control rats (Euclidean distance=152 ± 7 and 186 ± 8 in pre-training and post-training respectively; t(15)=3.32, P<0.01). The neural classifier performance was significantly improved in trained rats with respect to consonant discrimination (trained=78±1%; untrained=73±1.5%; t(15)=2.67, P<0.05), as well as vowel discrimination (trained=62±1%; untrained=59±0.6%; t(15)=2.91, P<0.05). Our finding of enhanced speech evoked activity after speech training is in agreement with human studies reporting significant increase in auditory evoked activity after perceptual training (Alain et al., 2007; Kraus et al., 1995).

In contrast to increased firing rates and increased neural differences observed in control rats after speech training, there was no significant increase of firing rate between trained and untrained rats within the pulsed-noise-reared (untrained=38±5, trained=33±4; t(12)=0.8,P=0.42) and speech-reared rats (untrained=35±4, trained=29±2; t(10)=1.6,P=0.14) (Figure 7). As a consequence there was no significant enhancement in neural discrimination of consonants or vowels, between trained and untrained counterparts of pulsed-noise-reared (average neural discrimination of consonant pairs Dad/Bad, Dad/Gad, Dad/Tad, Sad/Shad, Sad/Had, and Sad/Dad: untrained=74±1.1%, trained=74±2%, t(12)=0.13, P=0.90; average neural discrimination of vowel pairs Dad/Dud, Dad/Did, Dad/Dayed, Sad/Seed, Sad/Sead, and Sad/Sood: untrained=59±0.9%, trained=59±1.0, t(12)=0.18, P=0.86) and speech-reared rats(average neural discrimination of consonant pairs Dad/Bad, Dad/Gad, Dad/Tad, Sad/Shad, Sad/Had, and Sad/Dad: untrained=75±2.3%, trained=73±1.5%, t(10)=0.59, P=0.57; average neural discrimination of vowel pairs Dad/Dud, Dad/Did, Dad/Dayed, Sad/Seed, Sad/Sead, and Sad/Sood: untrained=61±1.6%, trained=60±0.6, t(10)=1.3, P=0.21). These results indicate that early acoustic experience of pulsed-noise or passive speech prevented the response strength enhancement effect of speech training that was observed in control rats. Collectively, our results demonstrate that the degree of learning induced plasticity that takes place in sensory cortices depends on previous sensory experience.

Training to discriminate complex stimuli induces changes in firing patterns of cortical neurons that are generalized to stimuli that share similar spectro-temporal characteristics. Speech training in human subjects shows equally significant MMN responses and increased neural synchrony for trained speech sounds and untrained speech sounds (Kraus et al., 1995; Tremblay et al., 2001). A1 neurons have been shown to generate similar pattern of activity for both trained and untrained marmoset twitter calls in ferrets (Schnupp et al., 2006). Consistent with these observations we observed similar A1 firing activity for both trained and untrained speech stimuli in our data. For example the peak firing rate within the first 100 ms of stimulus onset for 'sad' in control rats trained on 'sad' as their target stimulus was not significantly different from those trained on 'dad' as their target stimulus (t(6)=0.6951, P=0.51). As a consequence, there was no significant difference between the Euclidean distance patterns and the neural classifier performance between 'dad' trained rats and 'sad' trained rats. For example the classifier performance of consonant pairs of /D/ vs /T/, /D/ vs /B/, and /D/ vs /G/, was not different between control rats who were trained with /D/ as their target stimulus (79% ± 3) and control rats who were trained with /S/ (77% ± 2) as their target stimulus (t(6)=0.4, P=0.71). This observation was consistent across the pulsed-noise-reared rats (71% ± 2, 74% ± 3.5, for /D/ and /S/ target training respectively, t(6)=1.06, P=0.33) and speech-reared rats (73% ± 2, 74% ± 3, for /D/ and /S/ target training respectively, (6)=0.43, P=0.68). These observations suggest that enhanced A1 response

strength of speech-evoked activity is not specific to the learned tasks and hence may not be a necessary neural substrate for the perceptual learning ability.

4. DISCUSSION

In this study we tested the hypothesis that changes in spectral and temporal selectivity induced by early exposure to pulsed-noise or speech sounds were severe enough to produce significant changes in neural and behavioral discrimination of speech in later life. We found that the altered spectral and temporal characteristics of A1 neurons caused by pulsed-noise exposure did not significantly impair neural or behavioral discrimination of speech. We also found that the changes in temporal characteristics of neurons caused by speech exposure neither impaired nor improved the neural and behavioral discrimination of speech. Collectively, our results indicate that discrimination of speech is resistant to the alterations of spectral and temporal tuning properties of A1 neurons.

Consistent with previous studies (Insanally et al., 2010; Zhang et al., 2002; Zhou et al., 2008; Zhou et al., 2009) our results showed that early auditory experience of pulsed noise significantly alters the temporal modulation functions, response synchronization, tonotopic organization and frequency bandwidths of A1 neurons. The current study demonstrates that neural and behavioral discrimination of speech is resistant to the moderate degree of response degradation caused by pulsed-noise rearing. This finding complements the long list of experimental manipulations that highlight the surprisingly resilient nature of speech perception. Psychophysical experiments in both humans and animals have shown that speech recognition is highly resistant to spectral and temporal degradation of the speech signal as well as to masking noise as loud as the speech signal itself (Drullman et al., 1994; Miller et al., 1955; Ranasinghe et al., 2010; Shannon et al., 1995; Shetake et al., 2011 - in press). Even cochlear implants with a small number of spectral bands can restore speech perception of deaf individuals to near normal levels (Fishman et al., 1997). Our study demonstrates that neural discriminability of speech sounds presented in quiet remains unimpaired, despite a significant proportion of auditory neurons showing altered spectral and temporal tuning characteristics. These results suggest that studying neural responses to complex sounds would provide better insights of how auditory neurons process behaviorally relevant complex signals, rather than predictions based on simple tuning properties of auditory neurons. The current results however do not exclude the possibility that such altered response characteristics would impair the ability to process acoustically degraded speech sounds. Further studies are needed to determine whether these impairments cause speech processing to be more sensitive to signal degradation or background noise.

Our results suggest that early exposure to speech sounds presented passively does not provide an added advantage on behavioral and neural discrimination of speech. This observation is consistent with human psychophysical evidence showing that lexical context of speech plays a key role in early language development. Studies have shown that the amount of positive speech directed towards children (rather than total number of sounds heard) is best correlated with later vocabulary performance in school (Hoff, 2003; Weizman et al., 2001). Our results also agree with studies showing that exposure of human infants to foreign languages only leads to effective learning when the speech occurs in the context of a social interaction (Kuhl et al., 2003). Collectively, these observations suggest that plasticity of phonetic learning may require facilitation beyond passive auditory inputs, even during early development.

Learning induced plasticity in sensory cortices including audition, vision and somatosensation has been repeatedly described in both animals and humans (Bieszczad et al., 2010; Conner et al., 2005; Irvine et al., 1996; Recanzone et al., 1992a; Recanzone et al.,

1992b; Rutkowski et al., 2005). Despite several decades of research in this field, our understanding of the extent to which the direct behavioral consequences are associated with plasticity observed in primary sensory cortices, remains far from complete. While a number of studies provide evidence of stimulus specific plastic changes directly related to perceptual learning (Polley et al., 2006; Recanzone et al., 1993; Rutkowski et al., 2005), an equally strong amount of evidence suggests that such plasticity might not be a necessary condition for learned behavior (Brown et al., 2004; Scheich et al., 1993; Talwar et al., 2001; Yotsumoto et al., 2008). A recent hypothesis addressing these apparently contradictory observations suggested that plasticity of sensory cortices is a means of generating a large and diverse set of neurons that provide an efficient system to select the best circuitry, after which the neuronal properties renormalize (Reed et al., 2011). It is possible that early exposure and training caused significant but transient changes in speech sound responses.

The current results indicate that acoustic experience during early development can prevent the training-induced increase in speech-evoked activity at least at the time point evaluated. This finding suggests that the degree of learning associated neural plasticity generated in A1 can be significantly influenced by prior sensory experience, and is not simply driven by learning a particular stimulus configuration. Given that all three groups of rats (pulsed-noise-reared, speech-reared, and control) achieved the same degree of behavioral accuracy, it is unlikely that the neural plasticity that was only observed in control rats was the underlying neural substrate of the perceptual learning process. These results support the earlier hypothesis that large-scale changes in primary auditory cortex are not necessary for many forms of auditory learning (Brown et al., 2004; Scheich et al., 1993; Talwar et al., 2001; Yotsumoto et al., 2008)(Reed et al., 2011).

In summary, the current study demonstrates that the degree of plasticity generated by listening to a pulsed-noise or to passive speech during the critical period of auditory development does not impair or improve neural or behavioral discrimination of speech in later life. These results extend earlier observations that speech perception is robust by showing that speech perception is unimpaired by modest levels of neural degradation.

Acknowledgments

References

Alain C, Snyder JS, He Y, Reinke KS. Changes in auditory cortex parallel rapid perceptual learning. Cereb Cortex. 2007; 17:1074–84. [PubMed: 16754653]

Bieszczad KM, Weinberger NM. Representational gain in cortical area underlies increase of memory strength. Proc Natl Acad Sci U S A. 2010; 107:3793–8. [PubMed: 20133679]

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]

Chang EF, Bao S, Imaizumi K, Schreiner CE, Merzenich MM. Development of spectral and temporal response selectivity in the auditory cortex. Proc Natl Acad Sci U S A. 2005; 102:16460–5. [PubMed: 16263924]

Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT. Categorical speech representation in human superior temporal gyrus. Nat Neurosci. 2010; 13:1428–32. [PubMed: 20890293]

Cheour M, Ceponiene R, Lehtokoski A, Luuk A, Allik J, Alho K, Naatanen R. Development of language-specific phoneme representations in the infant brain. Nat Neurosci. 1998; 1:351–3. [PubMed: 10196522]

Conner JM, Chiba AA, Tuszynski MH. The basal forebrain cholinergic system is essential for cortical plasticity and functional recovery following brain injury. Neuron. 2005; 46:173–9. [PubMed: 15848797]

de Villers-Sidani E, Chang EF, Bao S, Merzenich MM. Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat. J Neurosci. 2007; 27:180–9. [PubMed: 17202485]

Delgutte B. Speech coding in the auditory nerve: II. Processing schemes for vowel-like sounds. J Acoust Soc Am. 1984; 75:879–86. [PubMed: 6707317]

Delgutte B, Kiang NY. Speech coding in the auditory nerve: I. Vowel-like sounds. J Acoust Soc Am. 1984; 75:866–78. [PubMed: 6707316]

Drullman R, Festen JM, Plomp R. Effect of temporal envelope smearing on speech reception. J Acoust Soc Am. 1994; 95:1053–64. [PubMed: 8132899]

Engineer, CT. Speech sound coding and training-induced plasticiity in primary auditory cortex. University of Texas at Dallas; Dallas: 2008.

Engineer CT, Perez CA, Chen YH, Carraway RS, Reed AC, Shetake JA, Jakkamsetti V, Chang KQ, Kilgard MP. Cortical activity patterns predict speech discrimination ability. Nat Neurosci. 2008; 11:603–8. [PubMed: 18425123]

Fishman KE, Shannon RV, Slattery WH. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. J Speech Lang Hear R. 1997; 40:1201–1215.

Foffani G, Moxon KA. PSTH-based classification of sensory stimuli using ensembles of single neurons. J Neurosci Meth. 2004; 135:107–120.

Green, DM.; Swets, JA. Signal detection theory and psychophysics. Peninsula Pub; 1989.

Han YK, Kover H, Insanally MN, Semerdjian JH, Bao S. Early experience impairs perceptual discrimination. Nat Neurosci. 2007; 10:1191–7. [PubMed: 17660815]

Hoff E. The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. Child Dev. 2003; 74:1368–78. [PubMed: 14552403]

Insanally MN, Albanna BF, Bao S. Pulsed noise experience disrupts complex sound representations. J Neurophysiol. 2010; 103:2611–7. [PubMed: 20200123]

Insanally MN, Kover H, Kim H, Bao S. Feature-dependent sensitive periods in the development of complex sound representation. J Neurosci. 2009; 29:5456–62. [PubMed: 19403813]

Irvine DR, Rajan R. Injury- and use-related plasticity in the primary sensory cortex of adult mammals: possible relationship to perceptual learning. Clin Exp Pharmacol Physiol. 1996; 23:939–47. [PubMed: 8911738]

Kawahara H. Speech representation and transformation using adaptive interpolation of weighted spectrum. IEEE Trans Acoustics Speech Signal Processing. 1997; 2:1303–1306.

Kilgard MP, Merzenich MM. Cortical map reorganization enabled by nucleus basalis activity. Science. 1998a; 279:1714–8. [PubMed: 9497289]

Kilgard MP, Merzenich MM. Plasticity of temporal information processing in the primary auditory cortex. Nat Neurosci. 1998b; 1:727–31. [PubMed: 10196590]

Kim H, Bao S. Selective increase in representations of sounds repeated at an ethological rate. J Neurosci. 2009; 29:5163–9. [PubMed: 19386912]

Kraus N, McGee T, Carrel TD, King C, Tremblay K, Nicol T. Central auditory system plasticity associated with speech discrimination training. Journal of cognitive neuroscience. 1995; 7:25–32.

Kuhl PK. Early language acquisition: cracking the speech code. Nat Rev Neurosci. 2004; 5:831–43. [PubMed: 15496861]

Kuhl PK. Brain mechanisms in early language acquisition. Neuron. 2010; 67:713–27. [PubMed: 20826304]

Kuhl PK, Tsao FM, Liu HM. Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. Proc Natl Acad Sci U S A. 2003; 100:9096–101. [PubMed: 12861072]

Kuhl PK, Kiritani S, Deguchi T, Hayashi A. Effects of language experience on speech perception: American and Japanese infants' perception of /ra/ and /la/. J Acoust Soc Am. 1997a; 102:3135–3136.

Kuhl PK, Andruski JE, Chistovich IA, Chistovich LA, Kozhevnikova EV, Ryskina VL, Stolyarova EI, Sundberg U, Lacerda F. Cross-language analysis of phonetic units in language addressed to infants. Science. 1997b; 277:684–6. [PubMed: 9235890]

Miller GA, Nicely P. An Analysis of Perceptual Confusions among some English Consonants. J Acoust Soc Am. 1955; 27:338–352.

Naatanen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio M, Alku P, Ilmoniemi RJ, Luuk A, Allik J, Sinkkonen J, Alho K. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature. 1997; 385:432–4. [PubMed: 9009189]

Nicholas JG, Geers AE. Effects of early auditory experience on the spoken language of deaf children at 3 years of age. Ear Hear. 2006; 27:286–98. [PubMed: 16672797]

Ohl FW, Scheich H. Orderly cortical representation of vowels based on formant interaction. Proc Natl Acad Sci U S A. 1997; 94:9440–9444. [PubMed: 9256501]

Pandya PK, Rathbun DL, Moucha R, Engineer ND, Kilgard MP. Spectral and temporal processing in rat posterior auditory cortex. Cereb Cortex. 2008; 18:301–14. [PubMed: 17615251]

Perez CA, Engineer CT, Jakkamsetti V, Carraway RS, Perry MS, Kilgard MP. Different Time Scales for the Neural Coding of Consonant and Vowel Sounds. Cereb Cortex. 2012 in press.

Pisoni DB. Auditory and phonetic memory codes in the discrimination of consonants and vowels. Perception and Psychophysics. 1973; 13:253–260.

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]

Puckett AC, Pandya PK, Moucha R, Dai W, Kilgard MP. Plasticity in the rat posterior auditory field following nucleus basalis stimulation. J Neurophysiol. 2007; 98:253–65. [PubMed: 17460101]

Ranasinghe, KG.; Vrana, WA.; Matney, CJ.; Mettalach, GL.; Rosenthal, TR.; Renfroe, EM.; Jasti, TK.; Kilgard, MP. Discrimination of degraded speech sounds by rats - behavior and physiology. ARO; Anaheim, CA: 2010.

Recanzone GH, Merzenich MM, Jenkins WM. Frequency discrimination training engaging a restricted skin surface results in an emergence of a cutaneous response zone in cortical area 3a. J Neurophysiol. 1992a; 67:1057–70. [PubMed: 1597697]

Recanzone GH, Merzenich MM, Schreiner CE. Changes in the distributed temporal response properties of SI cortical neurons reflect improvements in performance on a temporally based tactile discrimination task. J Neurophysiol. 1992b; 67:1071–91. [PubMed: 1597698]

Recanzone GH, Schreiner CE, Merzenich MM. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J Neurosci. 1993; 13:87–103. [PubMed: 8423485]

Reed A, Riley J, Carraway R, Carrasco A, Perez C, Jakkamsetti V, Kilgard MP. Cortical map plasticity improves learning but is not necessary for improved performance. Neuron. 2011; 70:121–31. [PubMed: 21482361]

Rutkowski RG, Weinberger NM. Encoding of learned importance of sound by magnitude of representational area in primary auditory cortex. Proc Natl Acad Sci U S A. 2005; 102:13664–9. [PubMed: 16174754]

Sachs MB, Young ED. Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. The Journal of the Acoustical Society of America. 1979; 66:470–9. [PubMed: 512208]

Scheich H, Simonis C, Ohl F, Tillein J, Thomas H. Functional organization and learning-related plasticity in auditory cortex of the Mongolian gerbil. Prog Brain Res. 1993; 97:135–43. [PubMed: 8234740]

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]

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. Science. 1995; 270:303–4. [PubMed: 7569981]

Shetake JA, Wolf JT, Cheung RJ, Engineer CT, Ram SK, Kilgard MP. Cortical activity patterns predict robust speech discrimination ability in noise. European Journal of Neuroscience. 2011 in press.

Svirsky MA, Teoh SW, Neuburger H. Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implantation. Audiol Neurootol. 2004; 9:224–33. [PubMed: 15205550]

Talwar SK, Gerstein GL. Reorganization in awake rat auditory cortex by local microstimulation and its effect on frequency-discrimination behavior. J Neurophysiol. 2001; 86:1555–72. [PubMed: 11600620]

Tremblay K, Kraus N, McGee T, Ponton C, Otis B. Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. Ear Hear. 2001; 22:79–90. [PubMed: 11324846]

van Wassenhove V, Nagarajan SS. Auditory cortical plasticity in learning to discriminate modulation rate. J Neurosci. 2007; 27:2663–72. [PubMed: 17344404]

Weizman ZO, Snow CE. Lexical input as related to children's vocabulary acquisition: effects of sophisticated exposure and support for meaning. Dev Psychol. 2001; 37:265–79. [PubMed: 11269394]

Yotsumoto Y, Watanabe T, Sasaki Y. Different dynamics of performance and brain activation in the time course of perceptual learning. Neuron. 2008; 57:827–33. [PubMed: 18367084]

Zhang LI, Bao S, Merzenich MM. Persistent and specific influences of early acoustic environments on primary auditory cortex. Nat Neurosci. 2001; 4:1123–30. [PubMed: 11687817]

Zhang LI, Bao S, Merzenich MM. Disruption of primary auditory cortex by synchronous auditory inputs during a critical period. Proc Natl Acad Sci U S A. 2002; 99:2309–14. [PubMed: 11842227]

Zhou X, Merzenich MM. Intensive training in adults refines A1 representations degraded in an early postnatal critical period. Proc Natl Acad Sci U S A. 2007; 104:15935–40. [PubMed: 17895375]

Zhou X, Merzenich MM. Enduring effects of early structured noise exposure on temporal modulation in the primary auditory cortex. Proc Natl Acad Sci U S A. 2008; 105:4423–8. [PubMed: 18332439]

Zhou X, Merzenich MM. Developmentally degraded cortical temporal processing restored by training. Nat Neurosci. 2009; 12:26–8. [PubMed: 19079250]

List of abbreviations

A1	primary auditory cortex
BW20	bandwidth at 20dB threshold
CVC	consonant-vowel-consonant
SPL	sound pressure level

1. Pulsed-noise-rearing and speech-rearing alters auditory cortex responses.

2. Early exposure does not alter speech discrimination ability.

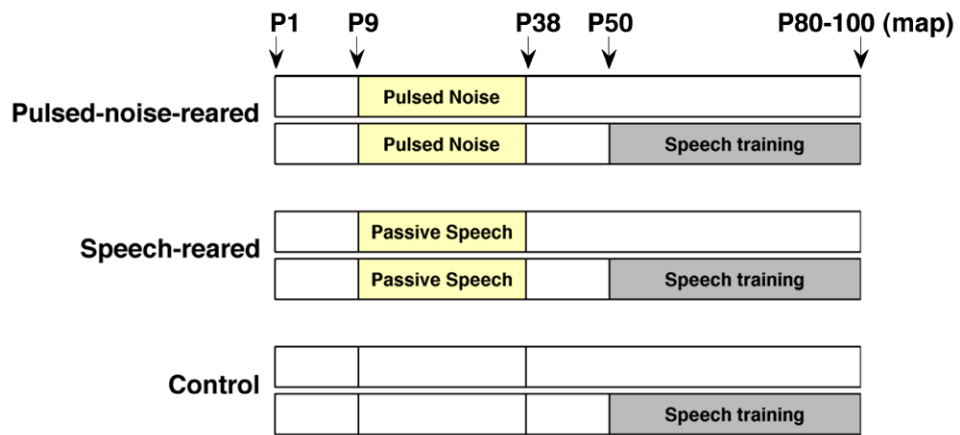3. Auditory cortex responses can account for robust speech discrimination.

Figure 1. Experimental timeline
Pulsed-noise-reared and speech-reared rat pups were exposed to a pulsed-noise and speech sounds respectively, from postnatal day 9 (P9) to postnatal day 38 (P38). Control group consisted of rats reared in an identical soundproof exposure booth without pulsed-noise or speech inputs (nearly half of the control rats) and rats reared in standard housing room conditions. Each group consisted of two subgroups as speech-trained (n=12 in each) and untrained. The trained subgroups started speech training on postnatal day 50 (P50). Physiology recordings from both trained and untrained rats of each group were done during a window of P80 to P100.
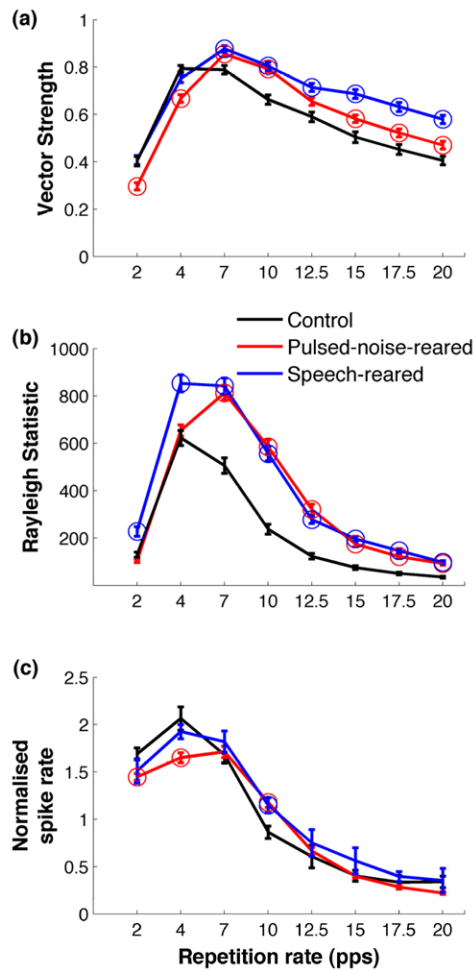
Figure 2. Post exposure temporal modulation functions
(a) Vector strengths measured at different repetition rates of a noise burst stimulus showed that pulsed-noise-rearing and speech-rearing result in an increased phase locking ability at higher repetition rates. (b) Rayleigh statistic showed significant phase locking (P<0.01) across all the repetition rates tested. Rayleigh statistic is a confidence measure for the vector strength where values above 13.8 indicate significant phase locking. (c) Normalized response strength of A1 neurons at different repetition rates of the noise burst stimulus. Normalized response was computed by the ratio of the average number of spikes evoked by the second through sixth bursts of the stimulus train to the number spikes evoked by the first burst. Exposure to pulsed noise produced increased A1 response strengths at 10 Hz and decreased response strengths at 2 and 4 Hz. Exposure to speech produced increased A1 response strengths at 10 Hz. Circles mark the points for pulsed-noise-reared and speech-reared rats which are significantly different from control rats using ANOVA and Dunnett's post hoc comparison (P<0.05). Error bars represent standard error of the mean for sites, control (n=129), pulsed-noise-reared (n=198), speech-reared (n=155). All data are from untrained rats.
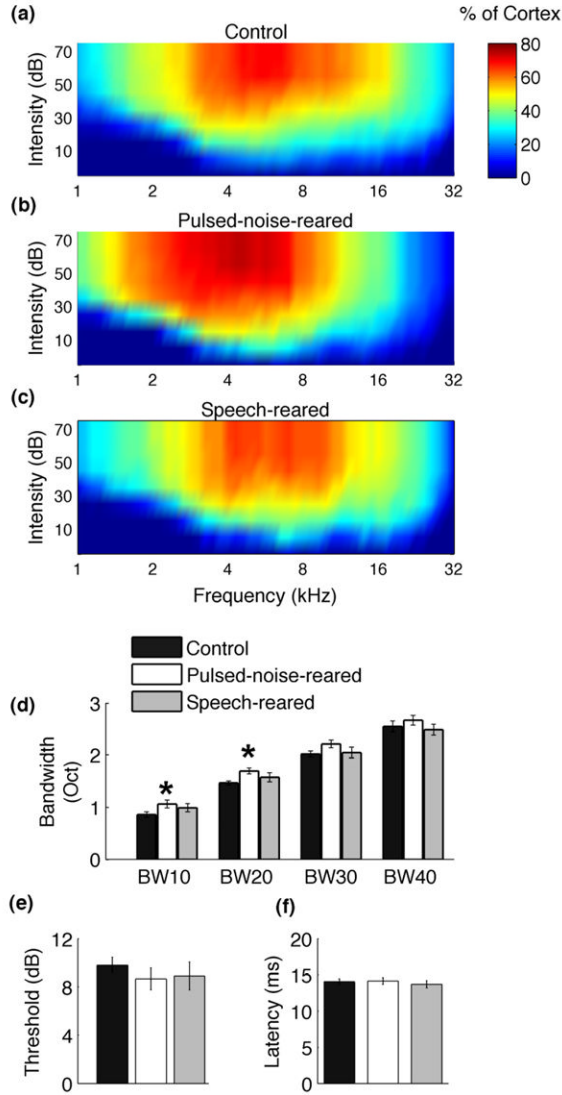
Figure 3. Changes in percent of cortex and tuning curve properties after pulsed-noise and speech exposure

Exposure to pulsed-noise resulted in an increased percent of cortex responding to low frequency tones, and decreased percent of cortex responding to high frequency tones. Color plots indicate the percentage of A1 neurons responded to a tone of a given frequency-intensity combination in control rats (a), pulsed-noise-reared rats (b), and speech-reared rats (c). Pulsed-noise-rearing resulted in a modest increase of tuning bandwidths while speech exposure did not change tuning bandwidths (d) BW10, BW20, BW30, BW40 stands for tuning bandwidths at 10 dB, 20dB, 30 dB and 40dB threshold levels respectively; stars indicate significant difference from control group (*=P<0.05) determined by one-way ANOVA for each BW level and Dunnett's one-tailed post hoc analysis. Both, pulsed-noise and speech exposure did not produce changes in response thresholds (e), or Latency (f) of A1 neurons. All data are from untrained rats. Error bars represent SE.
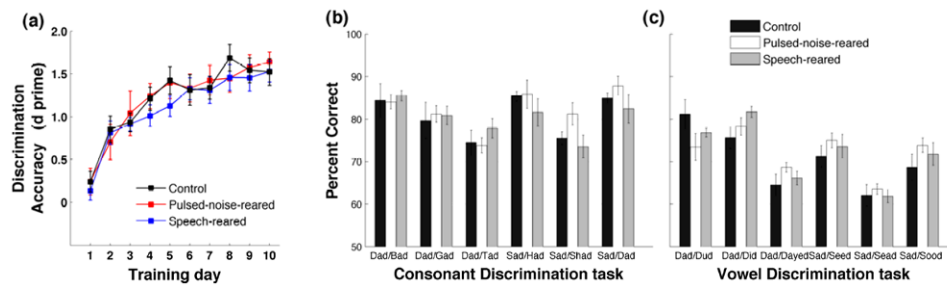
Figure 4. Behavioral discrimination of speech after pulsed-noise and speech exposure
Pulsed-noise-reared and speech-reared rats were not different from control rats in learning speech discrimination. d' measures the detection of target sound against all non-targets (six), computed for each rat, and averaged across groups. Percent correct discrimination of (b) consonant discrimination tasks and (c) vowel discrimination tasks for pulsed-noise-reared, speech-reared and control groups. Data are from the last four sessions of training. Error bars represent SE of the mean for rats.
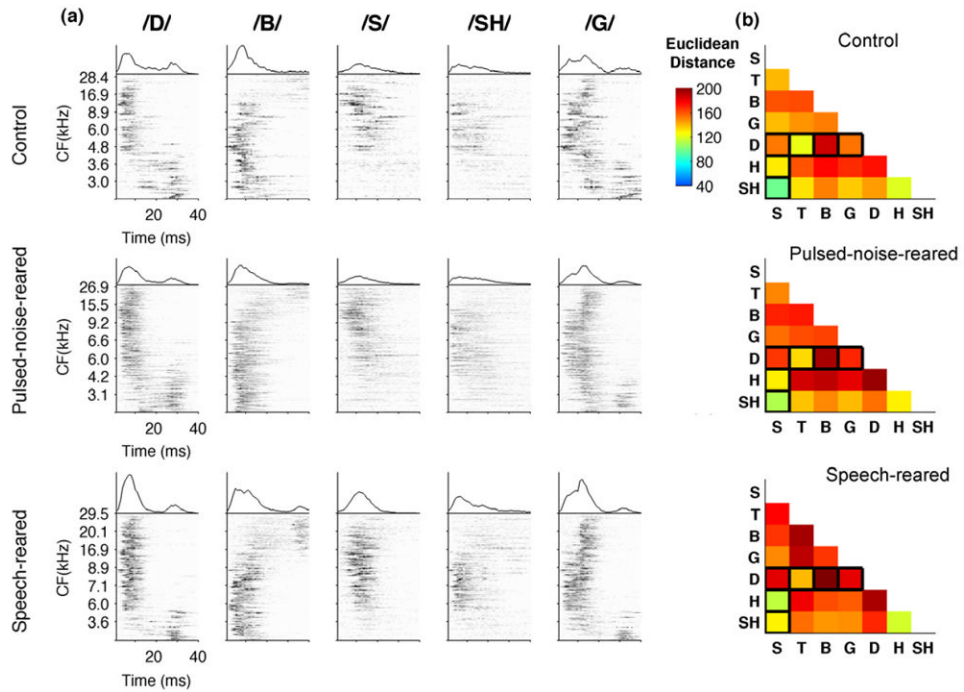
Figure 5. Neural representation of consonants in A1
(a) Neurograms depicting the onset response of A1 neurons to five consonants in control rats (top row), pulsed-noise-reared rats (middle row) and speech-reared rats (bottom row). Each neurogram plots the average poststimulus time histograms (PSTH) derived from 20 repeats, ordered by the characteristic frequency (kHz) of each recording site (y axis) of all the sites per group. Time is represented on the x axis (40 ms). The firing rate of each site is represented in gray-scale. For comparison, the mean population PSTH evoked by each sound is plotted above the corresponding neurogram. For better demonstration the neurograms are computed using ~ 150 sites recorded for each group. The Euclidean distance between the neurograms evoked by each consonant pair for control rats, n=8 (b), pulsed-noise-reared rats, n=6 and speech-reared rats, n=4. Euclidean distance was computed for each rat and averaged across the group. Distance measures were normalized by the most distinct pair where dark red squares indicate pairs that are highly distinct and blue indicates pairs that are less distinct. Pairs that were tested behaviorally are outlined in black. All data are from untrained rats.
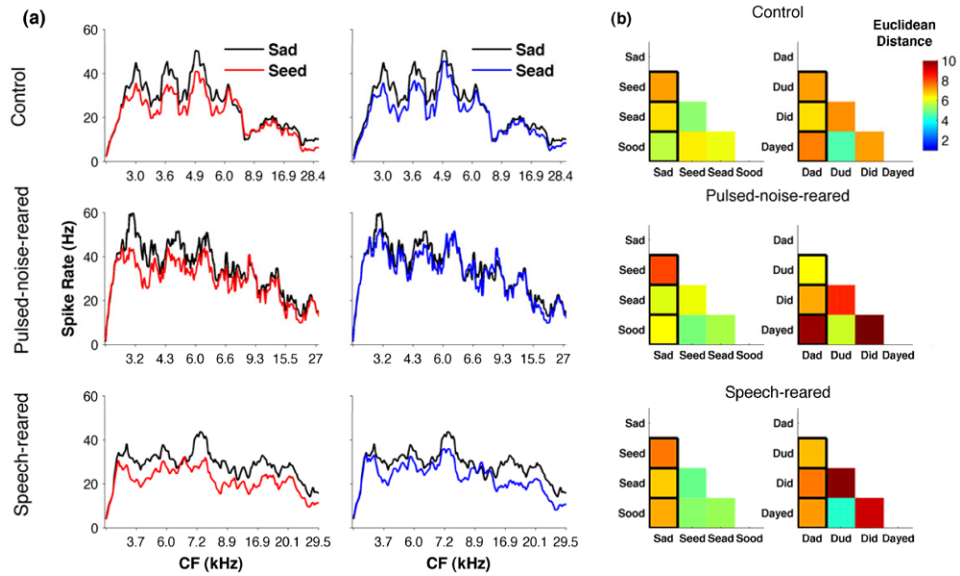
Figure 6. Neural representation of vowels in A1
(a) Spike count profiles depicting the response of A1 neurons to two vowel pairs in control rats (top row), pulsed-noise-reared rats (middle row) and speech-reared rats (bottom row). Each spike count profile plots the total number of spikes (averaged across 20 repeats) for the first 350 ms from the voicing onset, ordered by the characteristic frequency (kHz) of each recording site (x axis) of all the sites per group. For demonstration purposes number of spikes are smoothed across 10 adjacent sites. (b) The Euclidean distance between the spike count profiles evoked by each vowel pair for control rats (top row, n=8), pulsed-noise-reared (middle row, n=6), and speech-reared (bottom row, n=4). Euclidean distance was computed for each rat and averaged across the group. Distance measures were normalized by the most distinct pair where dark red squares indicate pairs that are highly distinct and blue indicates pairs that are less distinct. Pairs that were tested behaviorally are outlined in black. All data are from untrained rats.
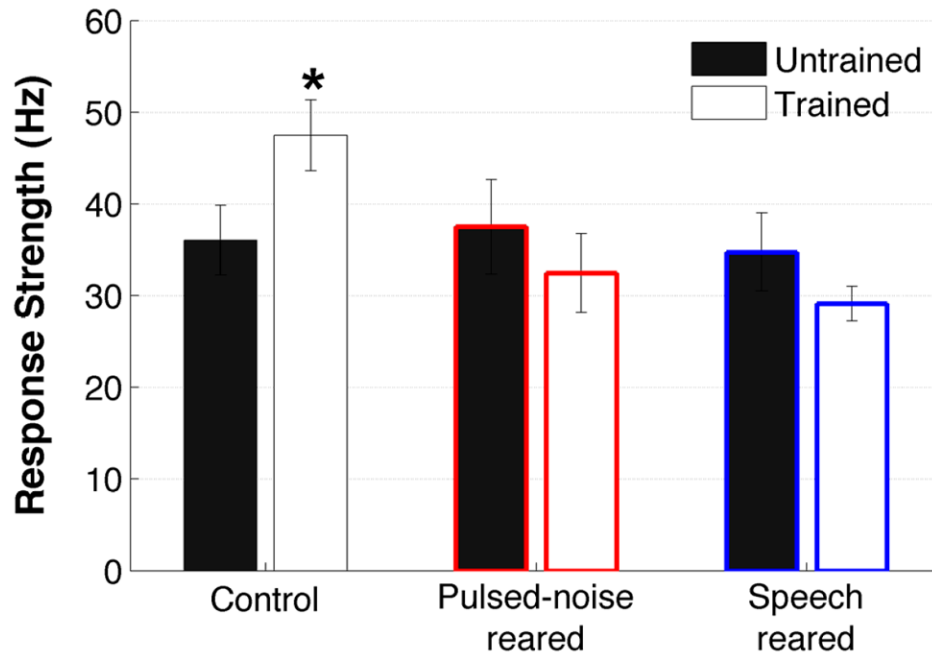
Figure 7. Speech training produced enhanced response strength in control group
The number of spikes fired during the first 100 ms of the stimulus onset by trained and
untrained counterparts of control rats control rats, pulsed-noise-reared and speech-reared
rats. Speech training increased the firing response in control rats but not in pulsed-noise-
reared and speech-reared rats. Stars indicate significant difference in trained groups
compared to untrained counterpart of each group. * =P<0.05. Error bars represent SE of the
mean for rats (n=8, n=8,control rats untrained and trained; n=6, n=8, pulsed-noise-reared
untrained and trained; n=4, n=8, speech-reared untrained and trained).