Published in final edited form as:

Neuroscience. 2013 November 12; 0: 80–97. doi:10.1016/j.neuroscience.2013.08.005.

Increasing diversity of neural responses to speech sounds across the central auditory pathway

Kamalini G. Ranasinghe^{1,2,3}, William A Vrana^{1,4}, Chanel J Matney^{1,5}, and Michael P. Kilgard^{1,2,6}

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

Abstract

Neurons at higher stations of each sensory system are responsive to feature combinations not present in lower levels. As a result, the activity of these neurons becomes less redundant than lower levels. We recorded responses to speech sounds from inferior colliculus and primary auditory cortex neurons of rats, and tested the hypothesis that primary auditory cortex neurons are more sensitive to combinations of multiple acoustic parameters compared to inferior colliculus neurons. We independently eliminated periodicity information, spectral information and temporal information in each consonant and vowel sound using a noise vocoder. This technique made it possible to test several key hypotheses about speech sound processing. Our results demonstrate that inferior colliculus responses are spatially arranged and primarily determined by the spectral energy and the fundamental frequency of speech, whereas primary auditory cortex neurons generate widely distributed responses to multiple acoustic parameters, and are not strongly influenced by the fundamental frequency of speech. We found no evidence that inferior colliculus or primary auditory cortex was specialized for speech features such as voice onset time or formants. The greater diversity of responses in primary auditory cortex compared to inferior colliculus may help explain how the auditory system can identify a wide range of speech sounds across a wide range of conditions without relying on any single acoustic cue.

Keywords

Multiple acoustic parameters; Neural response diversity; Noise vocoded speech; Redundancy reduction; Rat auditory system

INTRODUCTION

Neural activity patterns generated at higher stations of the sensory system are more diverse, hence less redundant, compared to lower stations (Friston, 2005). Adjacent primary auditory

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

²Corresponding author: The University of Texas at Dallas, School of Behavioral Brain Sciences, 800 West Campbell Road, GR41, Richardson, TX 75080-3021, Tel: 1-469-396-6098, Fax: 1-972-883-2491.

³kamalini@utdallas.edu

⁴wav031000@utdallas.edu

⁵cjmatney@gmail.com

⁶kilgard@utdallas.edu

The authors express no conflicts of interest.

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.

cortex (A1) neurons with similar frequency tuning generate distinct patterns in response to complex natural stimuli such as bird chirps, while inferior colliculus (IC) neurons with similar frequency tuning generate nearly identical patterns (Chechik *et al.*, 2006; Chechik & Nelken, 2012). A primary visual cortex neuron response gets modulated by the specific alignment of light intensity in the visual space, whereas neurons in the retina respond simply to the contrast at a particular point in space irrespective of its orientation (Hubel & Wiesel, 1962; 1968). Much of the compelling evidence of diverse firing activity in higher sensory stations comes from studies which identified neurons in sensory cortices that are selectively tuned to stimulus characteristics (i.e. feature detectors). Examples of these include the orientation selective neurons in the primary visual cortex (Hubel & Wiesel, 1962; 1968), face selective neurons in the inferotemporal cortex of the macaque (Gross *et al.*, 1972), auditory space specific neurons in the barn owl (Knudsen & Konishi, 1978), echolocation selective neurons in the bat auditory cortex (Suga, 1989), and birds' own song sensitive neurons in the auditory forebrains of zebra finches (Lewicki & Konishi, 1995; Lewicki, 1996). Other studies have identified that increased diversity arises as neurons at higher processing stations become sensitive to abstract feature combinations. For example, in A1, the redundancy was much reduced compared to IC because A1 neurons respond to many different acoustic parameters yet are less selective to stimulus identity than IC neurons (Chechik *et al.*, 2006). It has been proposed that both sparse coding and coarse coding strategies contribute to efficient representation of sensory information (Barlow, 1961).

It is not clear how the central auditory system represents human speech, which is perhaps the best studied complex acoustic signal (Fitch *et al.*, 1997). Acoustic analyses as well as theoretical studies have described distinct acoustic features for vowels and consonants defined by their spectral and temporal properties (Liberman *et al.*, 1967; Kent & Read, 2001). For example voiced stop consonants ('d', 'b', 'g') are defined by their longer voice onset time compared to unvoiced stop consonants ('p', 't', 'k'). Unique combinations of the first two formant frequencies (F1 & F2) and their trajectories can be used to define many of the English vowels (Peterson & Barney, 1952). Vowels are typically periodic with a relatively stable frequency distribution, while most consonants contain rapidly changing frequency patterns and noise bursts. Neurophysiological studies in animals have shown that auditory neurons can reliably encode the acoustic differences that distinguish different consonants and vowels (Steinschneider *et al.*, 1982; Steinschneider *et al.*, 1999; Steinschneider *et al.*, 2003; Wong & Schreiner, 2003; Steinschneider *et al.*, 2005; Engineer *et al.*, 2008; Shetake *et al.*, 2011; Perez *et al.*, 2012; Ranasinghe *et al.*, 2012a; Ranasinghe *et al.*, 2012b). The neural activity patterns generated in rat A1 and IC are significantly correlated with the behavioral discrimination of vowels and consonants in quiet (Engineer *et al.*, 2008; Perez *et al.*, 2012) as well as degraded conditions (Engineer *et al.*, 2008; Shetake *et al.*, 2011; Perez *et al.*, 2012; Ranasinghe *et al.*, 2012a; Ranasinghe *et al.*, 2012b). These results suggest that auditory processing convert the speech sounds into a spatiotemporal activity pattern that is sufficient to explain behavior. It remains unclear how the processing of complex speech like sounds change along the hierarchy of central auditory pathway.

In this study we examined the diversity of speech-evoked patterns across IC and A1. We used a noise vocoder to independently eliminate the periodicity information, spectral information, and temporal information in consonant and vowel sounds. This technique allowed us to record auditory responses of IC and A1 as aspects of the acoustic structure of speech is systematically removed. The aim of our study was to determine whether A1 neurons generate more diverse patterns to speech sounds compared to IC. We further explored whether A1 neurons are sensitive to speech features that IC neurons are not sensitive to, and whether A1 neurons are more sensitive to combinations of acoustic parameters compared to IC.

METHODS

Speech stimuli and noise vocoding of speech

We used seven English words including ‘dad’, ‘bad’, ‘sad’, ‘tad’, ‘dud’, ‘deed’, and ‘dood’, spoken by a native female English speaker in a double walled sound proof booth. The fundamental frequency and the spectral envelope of each recorded word were shifted up in frequency by a factor of two using the STRAIGHT vocoder (Kawahara, 1997) to better match the rat hearing range (Engineer *et al.*, 2008). For example the fundamental frequency of the octave shifted ‘Dad’ was 225 Hz. Each octave-shifted word was then degraded through a noise vocoder (Shannon *et al.*, 1995). The first step in the noise vocoder was to pass each word through a bank of bandpass filters. The number of bandpass filters used in this step varied between 2 and 16 (2, 4, 8, 16), and determined the four levels of spectral degradation (Figure 1). The overall bandwidth of the degraded speech sounds ranged from 300 to 11,000 Hz. Logarithmic filter spacing was used to set the bandwidths of the filters (Loizou *et al.*, 1999). Next, the temporal envelope of each analysis band was extracted by half-wave rectification and low-pass filtering. The low pass filter cutoff frequency varied between 4 Hz and 256 Hz (4, 16, 64, 256) determining the four levels of temporal degradation (Figure 1). The extracted envelope was then used to modulate a band of noise that was filtered with the same bandpass filters as the original speech. The modulated noise bands were summed to create the final degraded speech sound. Acoustic structures of degraded speech signals and sample audio material of processed speech has been published in a previous study (Ranasinghe *et al.*, 2012b). The signal processing was done using MatLab (MathWorks, Natick, MA). The four spectral levels and the four temporal levels created a combination of 7 degraded levels. In addition to the degraded speech sounds, we also used the undegraded versions of the seven words. A total of 56 speech stimuli (7 degraded levels $\times$ 7 words, plus 7 undegraded words) were used in this study.

Neural recording procedure

We recorded neural responses from 120 multi-unit sites of A1 from 5 rats and 118 multiunit sites of IC from 8 rats. All animals were experimentally naive, anesthetized, female Sprague-Dawley rats. Rats were anesthetized with pentobarbital (50 mg/kg), and received supplemental dilute pentobarbital (8 mg/ml) every ½ to 1 hour, subcutaneously, along with a 1:1 mixture of dextrose and Ringers lactate. The supplemental dosage was adjusted as needed to maintain areflexia. Heart rate and body temperature were monitored throughout the experiment. Auditory responses were recorded with Parylene-coated tungsten micro-electrodes (1–2 M Ω , FHC). For A1 recordings, four electrodes were lowered simultaneously to 600–700 μ m below the surface of the right cortex corresponding to layer V (Turner *et al.*, 2005). Each electrode penetration site was marked on a photograph of the surgically exposed temporal cortex, using cortical blood vessels as landmarks. As in earlier studies (Kilgard & Merzenich, 1998) Vornoni tessellation was used to transform the discretely sampled surface into a continuous map using the assumption that each point on the map has the response characteristics of the nearest recording site. A1 sites were identified based on latency and topography. For IC recordings, 1 or 2 Parylene-coated tungsten micro-electrodes (1–2 M Ω , FHC) were introduced through a hole located 9 mm posterior and 1.5 mm lateral to Bregma. Using a micromanipulator, electrodes were lowered at 200 μ m intervals along the dorsal ventral axis to a depth between 1,100 and 5,600 μ m. Blood vessels were used as landmarks as the electrode placement was changed in the caudal-rostral direction. As described in previous studies based on latency analysis and frequency topography, the IC recordings were identified to be from the central nucleus (Palombi & Caspary, 1996; Perez *et al.*, 2012).

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 curve of each site. Speech stimuli consisting of 49 degraded speech sounds and the seven undegraded speech sounds were presented after the tones. Each stimulus was played 20 times per site, in random order. The stimuli were presented every 2,300 ms at 65 dB SPL from a calibrated speaker mounted approximately 10 cm from the base of the left pinna. The intensity of all the speech sounds used in the study was adjusted so that the intensity of the loudest 100 ms was at 60 dB SPL. The recording techniques and data acquisition steps performed in the current study were identical to those published in earlier experiments (Engineer *et al.*, 2008; Perez *et al.*, 2012; Ranasinghe *et al.*, 2012b). Stimulus generation, data acquisition and spike sorting was performed with Tucker-Davis hardware (RP2.1 and RX5) and software (Brainware). Protocols and recording procedures were approved by the University of Texas at Dallas Institutional Animal Care and Use Committee.

Data Analysis

Previous studies reported that the number of spikes (spike count) generated over a longer time windows (i.e. 200 – 400 ms) is significantly correlated with representation of vowels, and the spike timing activity (1–10 ms precision) over shorter time windows (20–80 ms) is significantly correlated with representation of consonants, in the auditory system (Engineer *et al.*, 2008; Shetake *et al.*, 2011; Perez *et al.*, 2012; Ranasinghe *et al.*, 2012b). We analyzed the number of spikes (spike count profiles) generated by A1 and IC neurons as the vowel evoked response, and the spike timing activity (Post stimulus time histogram (PSTH) with 1 ms precision) of A1 and IC neurons as the consonant evoked activity. The analysis window for vowel sounds was 300 ms long beginning at the sound onset, and the analysis window for consonant sounds was 40 ms long beginning at the sound onset. We examined the spike count profiles of four vowels ('æ', ' ', 'i', and 'u' from the first 300 ms evoked activity of 'dad', 'dud', 'deed', and 'dood' respectively) and the spike timing patterns of four consonants ('d', 'b', 's', and 't' from the first 40 ms evoked activity of 'dad', 'bad', 'sad, and 'tad' respectively). Neural response at each site consisted of the average of 20 repetitions.

We used Pearson's correlation coefficient as a measure of neural diversity. Strong correlations between the activity recorded at different sites indicated less diversity and weak correlations indicated more diversity. Correlation of the vowel evoked response between two neurons was the Pearson's correlation coefficient between the number of spikes generated at each site in response to the four vowel stimuli ('æ', ' ', 'i', and 'u' from the 'dad', 'dud', 'deed', and 'dood' respectively). Correlation of the consonant evoked responses at two sites was the Pearson's correlation coefficient between the spike timing patterns (PSTH) generated by each neuron in response to the four consonants (the onset consonants 'd', 'b', 's', and 't' from 'dad', 'bad', 'sad, and 'tad' respectively).

The similarity of population response (A1 or IC) to each of the vowel pairs and consonant pairs was computed by Euclidean distance measure. We estimated the Euclidean distance between spike count profiles between vowel pairs. The Euclidean distance between two vowel pairs (i.e., of "dad" and "dud") was the square root of the sum of the squared differences between the number of spikes (300 ms window) for each recording site. The four vowels produced six vowel pair combinations ('dad vs dud', 'dad vs dood', 'dad vs deed', 'dud' vs 'dood', 'dud' vs 'deed', 'dood vs 'deed'). We estimated the Euclidean between two neurograms depicting the spike timing patterns evoked by two consonants. The Euclidean distance between a given neurogram pair is the square root of the sum of the squared differences between the firing rates for each bin (1 ms bins for 40 ms window) for each recording site. The four consonants produced six consonant pair combinations ('d' vs 'b', 'd' vs 's', 'd' vs 't', 'b' vs 's', 'b' vs 't', 's' vs 't').

To examine the relationship between the power spectra of each speech sound and the speech-evoked neuronal activity, we calculated Spearman's correlation coefficient between the acoustic energy for each sound contained within the frequency range of the tone frequency receptive field at 60 dB SPL and the number of spikes evoked by each sound. The stimulus set included the onset to four consonant sounds (i.e. 'd', 'b', 's' and 't') at five levels of spectral degradation (undegraded, 16, 8, 4, or 2 bands) for a total of twenty sounds. The onset response was simply the number of action potentials within the first 30 ms of sound onset. The spectral energy was measured as the first 20 ms following sound onset because, given the observed latencies, this interval was most likely to have generated the neural onset response. The filtering method used to extract the energy in each site's frequency receptive field was identical to that used to create the noise vocoded speech sounds. This analysis method provides only a coarse estimate of the expected response based on tone frequency tuning, but has the advantage of being simple and makes relatively few of complex assumptions about neural coding. All data analysis was done using MatLab (MathWorks, Natick, MA).

RESULTS

Speech Evoked Responses of IC and A1

We recorded speech evoked multiunit activity from 120 A1 sites and 118 IC sites (Table 1, tuning curve characteristics of A1 and IC sites), in response to four vowel sounds (æ, ' ', 'i', and 'u' from the 'dad', 'dud', 'deed', and 'dood' respectively) and four consonant sounds ('d', 'b', 's', 't'). As observed in earlier studies using awake and anesthetized preparations (Watanabe, 1978; Sinex & Chen, 2000; Cunningham *et al.*, 2002; Engineer *et al.*, 2008; Perez *et al.*, 2012), speech sounds evoked a brief phasic response for the initial consonant, and a sustained response for the vowel (Figure 2b for IC responses and Figure 2c for A1 responses). As expected from their more peripheral location in the auditory pathway, IC neurons responded to speech sounds earlier than A1 neurons (3.1 ± 1.2 ms earlier, $P < 0.05$, $t=5.8$). IC demonstrated a significantly higher speech evoked onset activity (mean firing rate = 162 ± 13 Hz and 30 ± 2 Hz for IC and A1 respectively; $P < 0.0001$, $t=10$) and a more pronounced sustained spike rate compared to A1 (Figure 2b & 2c). A1 evoked a sharp onset response for the initial consonant followed by a few peaks of sustained activity. These results are consistent with previous neurophysiological studies showing that the central stations of the auditory system emphasize transients over steady state acoustic properties of speech (Young, 2008).

Previous experiments reported that IC neural responses to bird vocalizations are highly redundant compared to A1 (Chechik *et al.*, 2002; Chechik *et al.*, 2006). Nearby IC neurons encode redundant information in the number of spikes (spike count) as well as the temporal patterns (spike timing) evoked by bird chirps. Given that both bird song and human speech share many parallels including complex acoustic characteristics of the signal and developmental plasticity driven by auditory learning (Doupe & Kuhl, 1999), it is reasonable to expect that ascending auditory neurons might generate similar patterns of diversity to speech sounds. We predicted that A1 neural responses to speech sounds would be more diverse (hence, less redundant) compared to IC. We examined the speech evoked activity generated in A1 and IC in response to four vowels (æ, ' ', 'i', and 'u' from the 'dad', 'dud', 'deed', and 'dood' respectively) and four consonants ('d', 'b', 's', 't').

A1 neural responses to vowels are more diverse than IC

It has been demonstrated in acoustic as well as behavioral experiments that the number of spikes (i.e. spike count) generated by auditory neurons is a significant predictor of vowel representation (Young, 2008; Perez *et al.*, 2012). We compared the number of spikes

generated by neurons at different recordings sites that were tuned to the same characteristic frequency. The spike count profiles captured the first 300 ms evoked activity of 'dad', 'dud', 'deed', and 'dood'. Examples of spike count measures are displayed for a pair of IC and a pair of A1 recording sites tuned to 1 kHz (Figure 3a and 3b) and for another two pairs tuned to 10 kHz (Figure 3e and f). The IC neurons tuned to the same characteristic frequency exhibited nearly identical spike count patterns. For example both members of the 1 kHz tuned IC pair generated relatively low number of spikes to 'dud' and more number of spikes to 'deed' (Figure 3a), and both members of the 10 kHz tuned IC pair generated increasing number of spikes in the order of 'dud', to 'deed', to 'dood' (Figure 3e). The strong and significant correlations between IC neurons (Figure 3c and 3g; $r=0.97$, $P<0.05$ and $r=0.97$, $P<0.05$ for 1kHz tuned pair and 10kHz tuned pair respectively) indicate that the second member of the IC pair would not add much to the information encoded by the first, demonstrating information redundancy. In contrast, A1 neurons tuned to the same characteristic frequency generated very different responses from each other. For example, the first A1 site tuned to 1 kHz generated relatively more spikes to 'dud' and few spikes to 'deed', while the other A1 site generated the opposite pattern (Figure 3b). Similarly, the first 10 kHz A1 site generated fewer spikes to 'dud' compared to 'dad', while the 2nd 10 kHz tuned A1 site generated the opposite pattern (Figure 3f). As a result correlation between A1 pairs were weak and failed to reach statistical significance (Figure 3d and 3h; $r=-0.55$, $P=0.45$ and $r=0.47$, $P=0.53$ for 1 kHz tuned pair and 10 kHz tuned pair respectively). The diverse firing patterns of similarly tuned sites suggests that A1 neurons encode independent information about each of the vowel stimuli. We quantified the correlation coefficients for all the pairs of IC and A1 recording sites with the same tuning for tones (Figure 3i and 3j). More than 70% of neuronal pairs generated correlation coefficients above 0.75 in IC (Figure 3i), whereas less than 20% of A1 pairs generated correlation coefficients above 0.75 (Figure 3j) ($P<0.0001$). These results demonstrate that similarly tuned IC neurons generate similar spike count profiles to vowels whereas similarly tuned A1 neurons generate different patterns, and confirm our initial prediction that A1 generates more diverse speech evoked patterns.

A1 neural responses to consonants are more diverse than IC

Psychophysical studies as well as neural and behavioral correlations in animals have demonstrated that consonants can be reliably distinguished based on acoustic information found within the first 40 ms of sound onset (Blumstein & Stevens, 1979; Jongman, 1989; Engineer *et al.*, 2008). We examined the precise temporal patterns (1 ms bin size) of neural activity during the first 40ms onset window for the consonants 'd', 'b', 's' and 't'. We compared the spike timing profiles of neurons tuned to the same best frequency in IC and A1. Examples of response profiles are displayed for a pair of IC and a pair of A1 sites tuned to 1 kHz (Figure 4a and 4d) and for another two pairs tuned to 10kHz (Figure 4b and 4e). The spike timing patterns of the IC sites tuned to the same best frequency were more similar than A1 sites tuned to the same best frequency. For example, 't' evoked a relatively low amplitude activity during the first 20 ms in both IC sites tuned to 1 kHz (Figure 4a). As a result the IC pairs tuned to the same frequency generated significant correlations (Figure 4a & 4b, $r=0.67$, $P<0.05$ and $r=0.54$, $P<0.05$, for 1kHz tuned pair and 10kHz tuned pair respectively). In contrast, the two A1 sites tuned to 1 kHz generated two different patterns for 't'. For example the 2nd A1 site showed a proportionately larger response during the last 20 ms that was absent in the 1st A1 sites' response (Figure 4d). The activity patterns generated by A1 neurons tuned to the same best frequency, hence, were not correlated to each other (Figure 4d & 4e, $r=0.13$, $P=0.10$ and $r=0.14$, $P=0.07$, for 1kHz tuned pair and 10kHz tuned pair respectively) These results demonstrated the greater diversity of A1 representations of consonants, compared to IC.

We estimated the correlation coefficients for all the IC and A1 pairs tuned to the same best frequency (Figure 4c and 4f, for IC and A1 respectively). 32% of IC pairs generated correlation coefficients above 0.7 (Figure 4c) while only 18% of A1 pairs generated correlation coefficients above 0.7 (Figure 4f) ($P < 0.01$). These results indicate that when encoding rapid acoustic transients, neurons tuned to the same best frequency in IC generate temporally coherent spikes, whereas similarly tuned A1 neurons fire relatively independently. Collectively, the increased diversity of A1 speech responses to both vowels and consonants suggests transformed sensitivity patterns across central auditory stations. Next, we examined the potential mechanisms that contribute to diverse tuning of A1 neurons.

IC and A1 responses are equally sensitive to spectral and temporal characteristics of speech sounds

In the ascending visual pathway, neurons become tuned to new features. For example, the neurons in primary visual cortex become tuned to specific orientations of the stimuli, whereas at the level of the visual thalamus neurons are not sensitive to stimulus orientation (Hubel & Wiesel, 1962; 1968). We examined whether A1 neurons are explicitly sensitive to speech features that IC neurons are not sensitive to, hence becoming more diverse in their neural response. We systematically degraded the spectral frequency details and temporal modulations of speech stimuli via a noise vocoder (see methods). The noise vocoder created four levels of spectrally degraded stimuli (number of channels = 16 bands, 8 bands, 4 bands and 2 bands), and four levels of temporally degraded stimuli (Low pass filter cut off = 256 Hz, 64 Hz, 16 Hz, 4 Hz). If A1 neurons are explicitly tuned to speech features, they would be expected to become less diverse as these features get eliminated with spectral and temporal degradation of the signal. For example, if a certain subset of A1 neurons are tuned to the longer voice onset time of 't' and do not fire for the shorter voice onset time of 'd', temporal degradation of the signal which eliminates this characteristic should make neural firing to 'd' and 't' more similar. Likewise, if a subset of A1 neurons are tuned to the combination of the first formant peak and the second formant peak (F1 and F2) of /æ/ (as in 'dad') and do not respond to the F1 and F2 combination of /i/ (as in 'deed'), spectral degradation which eliminates the F1 F2 resolution would make A1 responses more similar compared to undegraded stimuli. We examined the degree of neural similarity of IC and A1 at each stepwise reduction of spectral and temporal degradation of speech. As a measure of neural similarity we estimated the Euclidean distance between a given pair of evoked responses (for example between 't' and 'd' or 'dad' and 'deed'). Any set of neurons that are tuned to specific speech features such as formant transitions or voice onset time are bound to generate patterns that lie closer in Euclidean space (hence smaller Euclidean distances) with each stepwise degradation. We examined whether A1 would generate significantly smaller Euclidean distance differences compared to IC when speech sounds are systematically degraded. Specifically, we tested the hypothesis whether A1 neurons become tuned to acoustic features of speech that IC neurons are not sensitive to.

Consistent with previous reports (Perez *et al.*, 2012; Ranasinghe *et al.*, 2012b) the spike count profiles of the first 300 ms of the vowels generated unique patterns in auditory neurons (Figure 5). For example the undegraded 'dad' generated relatively higher number of spikes in low frequency tuned neurons compared to undegraded 'dood' (Figure 5a). These patterns were consistent in IC and A1, and also remained clearly distinct as speech sounds were spectrally degraded down to 8 bands (Figure 5a, 8 bands). With further degradation, for example when the spectral number of channels was limited to 2 bands, the patterns became less distinct (Figure 5a, 4 & 2 bands). We examined the Euclidean distance between each of the vowel pair combination (four vowels generated 6 pairs), at each spectral level (i.e. undegraded, 16 bands, 8 bands, 4 bands, 2 bands), and measured the fractional change

in Euclidean distance at each step. For example, we measured the Euclidean distance between vowel pairs at 16 bands as a percentage of Euclidean distance at undegraded stage (Figure 5b, 1st column in bar plot). In both A1 and IC the Euclidean distance between vowel pairs remained relatively high with 16 bands and 8 bands (Figure 5b). With more severe degradation (i.e. 2 bands), Euclidean distance decreased by a larger proportion. There was no difference between IC and A1 in the percent change in Euclidean distance at each step of degradation (Figure 5b, ANOVA, $F(1,5)=0.33$, $P=0.57$). This observation indicates that both A1 and IC neurons respond to spectral characteristics of vowels with similar degree of sensitivity. Our finding suggests that no special formant features are explicitly extracted from speech by A1 neurons that IC neurons are not sensitive to.

We examined the precise spike timing activity of the first 40 ms of speech-evoked activity for the four consonants. We created neurograms depicting precise spike timing information of 1 ms bins of each neuron arranged according to their characteristic frequency (Figure 6). As described in previous reports (Engineer *et al.*, 2008; Shetake *et al.*, 2011; Perez *et al.*, 2012; Ranasinghe *et al.*, 2012b), each consonant evoked a unique spatiotemporal activity pattern in auditory neurons. For example, 'd' evoked high frequency firing early in their response followed by low frequency tuned neurons. In contrast, 't' evoked an early high frequency tuned neuronal response and relatively less low frequency activity throughout (Figure 6a). These patterns were consistent in IC and in A1 and also remained clearly distinct as the speech sounds got low pass filtered with 256 Hz and 64 Hz (Figure 6a, 256 Hz). With further degradation, for example when the temporal modulations above 4 Hz are filtered, the patterns became less distinct (Figure 6, 4Hz). We examined the Euclidean distance between each of the consonant pair combination (four consonants generated 6 pairs), at each temporal degradation level (i.e. undegraded, 256 Hz, 64 Hz, 16 Hz, 4 Hz), and measured the fractional change of Euclidean distance at each step. For example, we measured the Euclidean distance between consonant neurograms at the 256 Hz stage as a percentage of Euclidean distance at the undegraded stage (Figure 6b, 1st column in bar plot). In both A1 and IC the Euclidean distance between consonant pairs remained relatively high at 256 Hz and 64 Hz degradations (Figure 6b), and showed a relatively larger drop with more severe degradation (i.e. 4 Hz). There was no significant difference in the percent change for IC and A1 (Figure 6b, ANOVA, $F(1,5)=0.13$, $P=0.72$). This observation indicates that both A1 and IC neurons are sensitive to similar spectral and temporal features of the consonants. Hence, it is unlikely that speech features defined by voicing onset or formant transitions alone are explicitly extracted by A1 neurons. Collectively, our results support the hypothesis that increased speech evoked neural diversity of A1 is not generated by neurons explicitly tuned to acoustic features of vowels or consonants. We predicted that A1 neurons are tuned to multiple acoustic parameters, while IC neurons are primarily driven by the presence of energy in their frequency receptive field.

Increased spectral and temporal information of vowel sounds make A1 responses more diverse and IC responses less diverse

To examine the diversity of neural responses in A1 and IC as we varied the spectral details and temporal modulations of speech, we measured the correlation coefficient between activity at different recordings sites that are tuned to a frequency within half an octave of each other's characteristic frequency (Figure 7a and 7b, open symbols; circles and squares corresponds to IC and A1 neural responses respectively, at each degradation stage). We also measured the correlation coefficients between neurons tuned to frequencies beyond half an octave of each other's characteristic frequency (Figure 7a and 7b, shaded symbols; circles and squares corresponds to IC and A1 neural responses respectively, at each degradation stage). With addition of spectral details as we increased the number of spectral bands from 4 to 16, A1 responses became more diverse. For example at 4 bands stage similarly tuned A1

neurons were significantly more correlated (i.e. less diverse) than neurons tuned to frequencies more than half an octave apart (Figure 7a, open and shaded squares at 4 bands), and with 8 and 16 spectral bands similarly tuned A1 neurons became as diverse as the neurons that are tuned to frequencies more than half an octave apart (Figure 7a, open and shaded squares at 8 bands and 16 bands). In contrast, IC neurons became more correlated, hence less diverse, in their firing as the spectral information of vowel sounds was increased from 4 bands to 16 bands (Figure 7a, open circles). Despite opposite patterns in their neural diversity, the response strength of IC and A1 remained similar across these spectral degradation levels (Figure 7c). The drop in A1 correlation at 2 bands compared to 4 bands (hence increased diversity) is likely due to the fact that at 2 bands the signal is dominated by noise rather than the spectral peaks resulting in recruitment of more neurons firing for noise. Since neural response to random noise would generate more diverse response patterns in neurons the overall neural activity would become less correlated. Collectively, these results demonstrate that, with increased spectral resolution of the speech signal A1 neurons became sensitive to additional acoustic information contained in the signal, whereas IC neurons became more spatially refined in their response. These findings suggest the prospect of A1 neurons being tuned to multiple acoustic parameters of speech, while IC neurons are primarily driven by the amount of energy that falls within in their frequency receptive field.

Temporal modulations of speech play a crucial role in human speech comprehension (Shannon *et al.*, 1995). Neurophysiological experiments in animals have reported that auditory neurons reliably encode temporal modulations of complex communication sounds including speech (Nelken, 2004; Narayan *et al.*, 2006; Perez *et al.*, 2012; Ranasinghe *et al.*, 2012b). We examined the diversity of spike count profiles evoked by vowels in IC and A1 as we systematically low pass filtered the temporal modulations of speech. Both A1 and IC neurons showed a similar degree of increased firing rate for vowels as we increased the temporal modulations (Figure 7d). This additional neural activity generated significantly more correlated patterns in similarly tuned IC neurons, compared to neurons that were located half an octave apart. For example, at 4 Hz the average correlation coefficient for similarly tuned IC neurons was 0.5 ± 0.02 , whereas at 64 Hz they became less diverse and the correlation coefficient was increased to 0.7 ± 0.02 (one-way-ANOVA, $F=20.12$, $P<0.001$; Tukey post hoc $P<0.05$) (Figure 7b, open circles). This finding indicates that IC neurons tuned to similar frequencies are also tuned to similar temporal modulations. In contrast, similarly tuned A1 neurons generated increasingly diverse patterns as temporal modulations were increased (Figure 7b, open squares). For example at 4 Hz, correlation coefficient of similarly tuned A1 sites was 0.3 ± 0.02 whereas at 64 Hz they became more diverse and the correlation coefficient dropped to 0.06 ± 0.02 (one-way-ANOVA, $f=15.23$, $P<0.001$; Tukey post hoc $P<0.05$) (Figure 7b, open squares). These results show that A1 neurons that have similar frequency tuning become sensitive to different components of temporal details.

IC neurons are more sensitive to periodicity than A1 neurons

Previous experiments have suggested that IC neurons are more sensitive to periodicity compared to A1 neurons. IC neurons have been shown to encode stimulus periodicity in the form of both spike rate and temporal patterns of discharge (Langner & Schreiner, 1988; Schreiner & Langner, 1988). Periodicity appears to be represented topographically in IC (Langner *et al.*, 2002). We observed that with the addition of pitch information IC neural activity became more diverse, whereas A1 activity was unaffected. For example compared to spectral degradation at 16 bands, which does not contain pitch information, similarly tuned IC neurons generated weaker correlations at undegraded stimuli that contain pitch information (Figure 7a, open circles at 16 bands and undegraded stage, one-way-ANOVA, $F=18.11$, $P<0.001$; Tukey post hoc, $P<0.05$). In contrast, there was no change in the degree of diversity in A1 (Figure 7a, open squares at 16 bands and undegraded stage, one-way-

ANOVA, $F=10.96$, $P<0.001$, Tukey post hoc, $P>0.05$). The fact that removing pitch reduces the number of evoked action potentials in IC but not A1 is consistent with the previous observation that IC neurons are more sensitive to periodicity than A1 neurons. This finding is also consistent with human imaging studies reporting that fMRI activity to periodic sounds is enhanced in the inferior colliculus (Griffiths *et al.*, 2001).

Potential differences in the firing coherence of IC and A1 can be viewed as a likely mechanism that would give rise to increased diversity of A1 compared to IC. For example if the A1 responses are less coherent (i.e. noisier), compared to IC responses, this variability will reduce the correlation coefficients of neurons tuned to similar characteristic frequency. We compared the neural diversity of IC and A1 using subsets of neurons with similar spike count variability. We estimated the coefficient of variation for the number of spikes generated by each recording site considering each individual presentation of speech sounds (i.e. 20 repetitions), and selected a subset of neurons that have identical coefficient of variation ($n=57$ and $n=104$ for IC and A1 respectively). Our analysis using the recording sites that have identical spike count variability confirmed our previous observation by showing that, with increased spectral and temporal details of speech A1 responses become more diverse and IC responses become less diverse (data not shown). The response strengths of the IC and A1 sites with identical spike count variability remained similar across the degradation levels (data not shown). The fact that the diversity patterns remained unchanged indicates that greater response diversity and not greater noise is responsible for the difference in correlations between similarly tuned IC and A1 neurons.

A1 neurons are sensitive to multiple acoustic parameters of speech

We hypothesized that the greater response diversity in A1 is primarily due to tuning for multiple acoustic parameters of speech, while IC neurons are primarily driven by the amount of energy that falls within in their frequency receptive field. We predicted that speech responses from pairs of A1 sites with similar frequency selectivity (bandwidth) or latency would be more correlated than pairs of A1 sites with less similar selectivity or latency. We expected that bandwidth and latency similarity would have less influence in IC because response similarity is dominated by characteristic frequency.

We analyzed the diversity patterns of similarly tuned (neurons tuned to characteristic frequencies within half an octave) A1 and IC neurons and examined the effect of similarity in the bandwidth at 10dB threshold (bandwidth10) and at the peak firing latency. As predicted, responses of sites with similar bandwidths (less than 1 octave difference) were more correlated than sites with different bandwidths (more than 1 octave difference). With increasing spectral and temporal information, nearby A1 neurons with similar bandwidths generated more correlated activity patterns than nearby A1 neurons with dissimilar bandwidths (Figure 8b, open and shaded squares indicate sites with similar bandwidths and sites with dissimilar bandwidths respectively). In contrast, IC activity was highly correlated regardless of whether the sites had similar or dissimilar frequency bandwidths (Figure 8a). For example, when temporal information is increased from 16 Hz to 256 Hz to undegraded there was no difference between the correlations generated in sites with similar and dissimilar bandwidths (Figure 8a, open and shaded circles along the Un, 256, 64 and 16 stages). A similar pattern was observed for pairs of sites with similar response latencies. With increasing spectral and temporal information in the signal, nearby A1 neurons with similar firing latencies (i.e. within 1 ms of each other's) generated more correlated activity patterns than nearby A1 neurons with dissimilar firing latencies (i.e. latency separated >1 ms) (Figure 8d, open squares show higher correlations than shaded squares showing sites with similar latencies and sites with dissimilar latencies respectively). In contrast, nearby IC neurons generated equally correlated activity patterns regardless of their response latency

characteristics (Figure 8c). These results support our initial hypothesis that speech evoked activity in A1 is further shaped by the fine structure of speech sounds compared to IC.

Increased spectral and temporal information of consonants make modest changes in diversity of IC and A1 responses

In contrast to spectral changes of vowels, consonants contain rapidly changing spectral information and noise bursts in their acoustics. To examine the diversity of consonant evoked neural responses as we changed the spectral and temporal content of speech, we measured the correlation coefficient of spike timing profiles (40 ms onset window, 1 ms precision) between sites that are tuned to a frequency within half an octave of each other's characteristic frequency (Figure 9a and 9b, open symbols; circles and squares corresponds to IC and A1 neural responses respectively, at each degradation stage). We also measured the correlation coefficients between neurons tuned to frequencies beyond half an octave of each other's characteristic frequency (Figure 9a and 9b, shaded symbols; circles and squares corresponds to IC and A1 neural responses respectively, at each degradation stage). Spectral degradation only demonstrated modest changes in neural diversity patterns in either IC or A1. For example, increasing spectral number of bands from 2 to 16 did not change the correlation coefficients or the response strengths, in either IC or in A1 (correlation coefficients = 0.42 ± 0.01 and 0.42 ± 0.01 for IC at 2 and 16 bands respectively, (one-way-ANOVA, $F=14.64$, $P<0.001$, Tukey post hoc $P>0.05$) & correlation coefficients = 0.41 ± 0.01 and 0.42 ± 0.01 for A1 at 2 and 16 bands respectively, (one-way-ANOVA, $F=9.1$, $P<0.001$, Tukey post hoc $P>0.05$)(Figure 9a). In contrast, decreasing temporal modulations dramatically reduced the response strength and correlation coefficients in both IC and A1. For example, reduction of the maximum envelope modulation from 64 Hz to 4 Hz reduced the normalized spike rate from 1.15 ± 0.1 to 0.20 ± 0.03 and 1.15 ± 0.05 to 0.20 ± 0.03 in A1 and IC respectively (Figure 9d). This was accompanied by a large drop in correlations between spike timing profiles of both nearby and distant neuronal pairs, in both IC and A1, which resulted from lack of neural firing (Figure 9b). These results are consistent with numerous earlier reports that suggest temporal modulations play a crucial role in representation of consonants in auditory system (Ahissar *et al.*, 2001; Engineer *et al.*, 2008; Huetz *et al.*, 2011; Shetake *et al.*, 2011).

Our findings of IC and A1 neural diversity patterns generated by consonants were replicated in a subset of IC and A1 neurons with identical spike count variability ($n=49$ and $n=93$ for IC and A1 respectively), demonstrating that greater response diversity of A1, and not greater noise is responsible for the difference in correlations between similarly tuned IC and A1 neurons. We found that tuning characteristics such as bandwidths and response latencies of nearby A1 neurons as well as nearby IC neurons significantly affect the spike timing patterns evoked by consonants. For example in both IC and in A1 (two-way-ANOVA, $F=0.35$, $P=0.56$ for interaction effect), nearby neurons that had similar bandwidth responses at 10 dB above threshold (i.e. within 1 octave) generated more correlated spike timing patterns than nearby neurons that had different bandwidth responses (i.e. differ by >1 octave)(Figure 10a and 10b). Similarly, in both IC and A1 (two-way-ANOVA, $F=0.16$, $P=0.70$ for interaction effect), nearby neurons with similar firing latencies (i.e. within 1 ms of each other's) generated more correlated activity patterns than nearby neurons with dissimilar firing latencies (i.e. beyond 1 ms of each other's)(Figure 10c and 10d). These results show that neurons in both IC and A1 are sensitive to fine details of acoustic structure of consonants. This finding is compatible with our previous observation that consonants generated relatively less degree of diversity along the central auditory pathway (i.e. from IC to A1) compared to vowels (Figure 3 & 4 data).

IC responses are dominated by the spectral energy of the speech signal

Based on our observation that IC neurons are primarily driven by the amount of energy that falls within their frequency receptive field, we hypothesized that the speech-evoked activity of IC neurons' would be predictable from their frequency-intensity receptive field. We calculated the correlation between each site's speech evoked activity and the acoustic energy in that site's receptive field (see methods for details). Although this simple method allows only a coarse estimate of the expected neural response, most IC sites (64%) exhibited a significant correlation between acoustic energy and neural response. Fewer A1 sites (38%) exhibited a significant correlation between spectral energy and neural response and the average correlation was significantly lower than IC ($r = 0.48 \pm 0.03$ and $r = 0.31 \pm 0.03$ for IC and A1 respectively; $t=4.6$, $P<0.0001$). This observation was consistent across a wide range of temporal analysis windows and whether or not the degraded versions of the speech sounds were included. These results support our earlier observation that A1 responses are more diverse and less well predicted by the amount of energy that falls within their frequency receptive field compared to IC responses.

In summary, our results demonstrate that A1 neurons generate more diverse speech evoked responses than IC neurons. Neurons with similar frequency tuning in IC generated more correlated responses compared to neurons with similar frequency tuning in A1. This observation was consistent across the neural responses generated for both consonants and vowels. We have also demonstrated that IC neurons and A1 neurons show similar degree of sensitivity to spectral and temporal properties of the speech signal. The degree of neural discriminability generated by spike rate profiles of vowel responses and spike timing profiles of consonant response were similar in A1 and IC. We also found that increased spectral and temporal information of vowel sounds made A1 responses more diverse while IC responses were less diverse. There was only a modest degree of diversity differences between IC and A1 responses to consonant sounds. Further analysis of neural responses according to each neurons' bandwidth and latency tuning properties showed that A1 neurons are more sensitive to multiple acoustic properties of speech sounds. IC neurons were more sensitive to periodic information compared to A1. The response of IC neurons was dominated by the spectral energy of the speech signal.

DISCUSSION

In this study, we recorded the responses of IC and A1 neurons to a set of consonants and vowels, and tested the hypothesis that A1 neurons generate more diverse patterns to speech sounds compared to IC neurons. We independently eliminated the periodicity information, spectral information and temporal information of speech using a noise vocoder, and examined the diversity of speech-evoked neural activity patterns at two stations of the ascending auditory pathway. We found that A1 neurons with similar frequency tuning generate distinct patterns in response to speech, while IC neurons with similar frequency tuning generate nearly identical patterns. Our results further suggest that IC responses are spatially organized and largely determined by the spectral energy and the fundamental frequency of speech sounds, whereas A1 responses are more sparsely distributed, are highly sensitive to combinations of spectral and temporal parameters, and are not influenced by the fundamental frequency.

The current results compliment previous findings of increasing diversity of neurons in the ascending sensory pathway and support the hypothesis that neurons at higher stations of processing become tuned to more complex and independent features (Barlow, 1961; Chechik *et al.*, 2002; Chechik *et al.*, 2006). Our results demonstrate that the reduced redundancy of A1 is generated by diverse speech evoked activity independent of greater noise in neural firing. We also show that A1 neurons with distinct tuning properties (i.e.

neurons with different bandwidths, latency etc) generate more distinct firing patterns compared to IC neurons of comparable tuning calibers. While our results support the hypothesis that greater diversity is contributed by auditory neurons at higher stations responding to multiple acoustic properties of complex sounds, we do not yet know what exact features of speech evoke such diversity in A1 neurons. Response characteristics such as whether the neurons are monotonic or non-monotonic or whether the neurons are EE (contralateral and ipsilateral excitatory) or EI (contralateral excitatory and ipsilateral inhibitory) are likely to contribute to response diversity. Future studies are needed to determine how monotonicity and binaurality affect the diversity of auditory responses.. Further experimentation using methods that allow systematic manipulation of specific acoustic features (eg: Klatt synthesizer) (Klatt, 1980), would be a fascinating approach to tap into the properties that contribute directly to increased response diversity in A1 compared to IC.

Since the first demonstration of behavioral speech discrimination by chinchilla, several studies have reported rats, cats, monkeys and birds ability to reliably discriminate vowels and consonants (Kuhl & Miller, 1975; Kluender *et al.*, 1987; Dooling *et al.*, 1989; Ramus *et al.*, 2000; Reed *et al.*, 2003; Wong & Schreiner, 2003; Porter *et al.*, 2011). Neurophysiological recordings in animals have shown that activity patterns evoked by speech sounds in auditory neurons are well correlated with behavioral discrimination of speech (Engineer *et al.*, 2008; Perez *et al.*, 2012). Recent experiments have shown that such neural correlates continue to exist even after the acoustic content of the signal is significantly degraded by masking noise or by eliminating spectral and temporal details of speech (Shetake *et al.*, 2011; Ranasinghe *et al.*, 2012b). These findings suggest that auditory processing of speech demands no more than general neural mechanisms shared across most mammals including humans. The current study supports this hypothesis by showing that speech evoked neural activity is effectively transformed across central auditory stations by neurons becoming sensitive to multiple acoustic features, rather than becoming explicit single feature detectors. The ability of auditory neurons to generate distinct sensory patterns based on spectral and temporal properties of complex signals provides evidence for potential neural mechanisms shared by humans and animals. The current study further supports the evidence that greater diversity of central auditory stations is not limited to animal's own vocalizations, but extends to a wider range of non-species specific complex sounds. It is likely that auditory processing mechanisms evolved to process spectral and temporal information of complex animal vocalizations – from bird song to frog calls to human speech - whose general acoustic properties are extensively shared.

Auditory systems conspicuously transform sensory information to guide behavioral responses from a wide variety of sounds ranging from pure tones to complex animal vocalizations. The adeptness of auditory function is demonstrated by the apparently seamless integration of inputs into dynamic category boundaries and the highly accurate perception of degraded auditory inputs (i.e. presence of masking noise or spectral and temporal degradation of the signal). Both humans and animals are able to discriminate speech sounds with high accuracy regardless of speaker variability or masking noise as loud as the speech signal itself (Miller & Nicely, 1955; Kuhl & Miller, 1975; Shetake *et al.*, 2011). The current study shows that A1 neurons are able to encode distinct characteristics of speech sounds more efficiently (i.e. less number of spikes), and independently, compared to IC. It is reasonable to expect that distributed representations of speech features by neurons sensitive to multiple acoustic parameters provide more resistance against signal degradation and more efficient integration into dynamic category boundaries, than neural representations extracting a single speech feature, such as voice onset time or formant transition. Furthermore, neural representations based on sensitivities to multiple acoustic parameters of complex sounds provide a logical explanation of how auditory neurons effectively transform

simple acoustics of pure tones as well as complex spectro-temporal patterns of human speech.

Sensory information is transformed via different stages of a processing hierarchy to generate useful information to guide behavioral outputs. From peripheral sensory organs to central processing stations the neural representations range from perfect copies of stimulus characteristics to extremely complex and context-dependent responses (Kilgard, 2012). In this study, we examined the transformation of acoustic information contained in speech at two key stages of the central auditory pathway. Our results suggest that, at more central stages of auditory pathway, neurons become sensitive to combinations of acoustic parameters hence generating more diverse firing patterns (Buonomano & Merzenich, 1995). However the degree of neural distinctiveness generated in spatiotemporal activity patterns evoked by consonants and spatial patterns evoked by vowels were comparable in IC and A1. For example the fractional change of Euclidean distance between vowel pairs when spectral information was filtered out was not different from IC to A1. This finding indicates that the ability to discriminate acoustic features of vowels and consonants is relatively high even at the level of IC (Chen *et al.*, 2012). It is likely that A1 is involved in a role that is beyond simple representation of acoustic patterns of speech. This observation is consistent with studies showing that functional specializations can exist within A1 (Middlebrooks & Zook, 1983; Schreiner & Mendelson, 1990; Read *et al.*, 2001) as well as models suggesting A1 can represent abstract stimulus characteristics that are less dependent on acoustic based feature representations (Nelken, 2004).

Future studies are needed to evaluate the potential effect of anesthesia and attention on cortical and subcortical processing of speech sounds. Earlier studies have reported basic aspects of neural responses to sound are comparable in awake and anesthetized preparations of auditory recordings (Watanabe, 1978; Steinschneider *et al.*, 1982; Sinex & Chen, 2000; Cunningham *et al.*, 2002; Engineer *et al.*, 2008), although important differences have also been described (Anderson *et al.*, 2006; Rennaker *et al.*, 2007). A1 responses recorded both under anesthesia and in awake rats have shown to be well correlated with behavioral discrimination of consonants (Engineer *et al.*, 2008). However, anesthesia can reduce the maximum rate of click trains an A1 neuron is able to respond (Anderson *et al.*, 2006; Rennaker *et al.*, 2007). Although human imaging studies have shown little effect of attention on primary auditory cortex responses to speech sounds, single unit studies are needed to better understand how attention might shape the cortical representation of speech (Grady *et al.*, 1997; Hugdahl *et al.*, 2003; Christensen *et al.*, 2008). Future research exploring the relative diversity of neural activity patterns in peripheral versus central stations in awake animals will offer insights into further understanding of speech sound processing in the auditory system.

Acknowledgments

The authors would like to thank T. R. Rosenthal, E. M. Renfro, T. K. Jasti, E. Tran, K. Ram, R. Cheung, H. Shepard, Z. Ghneim, C. Xie, D. Vuppula, T. Nguyen and R. Joseph for help with behavioral training. We gratefully acknowledge the comments and suggestions given by Israel Nelken and Andrew Sloan, on earlier versions of the manuscript. We would also like to thank late P.C. Loizou in his assistance in signal processing and H. Abdi for his guidance in statistical analyses. This work was supported by Award Numbers R01DC010433 and R15DC006624 from the National Institute on Deafness and Other Communication Disorders.

ABBREVIATIONS

A1	Primary auditory cortex
F1	First formant

F2	Second formant
IC	Inferior colliculus
PSTH	Post stimulus time histogram

References

- Ahissar E, Nagarajan S, Ahissar M, Protopapas A, Mahncke H, Merzenich MM. Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc Natl Acad Sci U S A*. 2001; 98:13367–13372. [PubMed: 11698688]
- Anderson SE, Kilgard MP, Sloan AM, Rennaker RL. Response to broadband repetitive stimuli in auditory cortex of the unanesthetized rat. *Hear Res*. 2006; 213:107–117. [PubMed: 16466874]
- Barlow, H. Possible principles underlying the transformation of sensory messages. In: Rosenblith, IW., editor. *Sensory communication*. MIT Press; 1961. p. 217-234.
- Blumstein SE, Stevens KN. Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. *J Acoust Soc Am*. 1979; 66:1001–1017. [PubMed: 512211]
- Buonomano DV, Merzenich MM. Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*. 1995; 267:1028–1030. [PubMed: 7863330]
- Chechik G, Anderson MJ, Bar-Yosef O, Young ED, Tishby N, Nelken I. Reduction of information redundancy in the ascending auditory pathway. *Neuron*. 2006; 51:359–368. [PubMed: 16880130]
- Chechik, G.; Globerson, A.; Anderson, M.; Young, ED.; Nelken, I.; Tishby, N. Group redundancy measures reveal redundancy reduction along the auditory pathway. TGD; Becker, S.; Ghahramani, Z., editors. MIT; Cambridge, MA: 2002. p. 173-180.
- Chechik G, Nelken I. Auditory abstraction from spectro-temporal features to coding auditory entities. *Proc Natl Acad Sci U S A*. 2012
- Chen C, Read HL, Escabi MA. Precise feature based time scales and frequency decorrelation lead to a sparse auditory code. *J Neurosci*. 2012; 32:8454–8468. [PubMed: 22723685]
- Christensen TA, Antonucci SM, Lockwood JL, Kittleson M, Plante E. Cortical and subcortical contributions to the attentive processing of speech. *Neuroreport*. 2008; 19:1101–1105. [PubMed: 18596608]
- Cunningham J, Nicol T, King C, Zecker SG, Kraus N. Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. *Hear Res*. 2002; 169:97–111. [PubMed: 12121743]
- Doolling RJ, Okanoya K, Brown SD. Speech perception by budgerigars (*Melopsittacus undulatus*): the voiced-voiceless distinction. *Percept Psychophys*. 1989; 46:65–71. [PubMed: 2755763]
- Doupe AJ, Kuhl PK. Birdsong and human speech: common themes and mechanisms. *Annu Rev Neurosci*. 1999; 22:567–631. [PubMed: 10202549]
- Engineer CT, Perez CA, Chen YH, Carraway RS, Reed AC, Shetake JA, Jakkamsetti V, Chang KQ, Kilgard MP. Cortical activity patterns predict speech discrimination ability. *Nat Neurosci*. 2008; 11:603–608. [PubMed: 18425123]
- Fitch RH, Miller S, Tallal P. Neurobiology of speech perception. *Annu Rev Neurosci*. 1997; 20:331–353. [PubMed: 9056717]
- Friston K. A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci*. 2005; 360:815–836. [PubMed: 15937014]
- Gawne TJ, Kjaer TW, Hertz JA, Richmond BJ. Adjacent visual cortical complex cells share about 20% of their stimulus-related information. *Cereb Cortex*. 1996; 6:482–489. [PubMed: 8670673]
- Grady CL, Van Meter JW, Maisog JM, Pietrini P, Krasuski J, Rauschecker JP. Attention-related modulation of activity in primary and secondary auditory cortex. *Neuroreport*. 1997; 8:2511–2516. [PubMed: 9261818]
- Griffiths TD, Uppenkamp S, Johnsrude I, Josephs O, Patterson RD. Encoding of the temporal regularity of sound in the human brainstem. *Nat Neurosci*. 2001; 4:633–637. [PubMed: 11369945]

- Gross CG, Rocha-Miranda CE, Bender DB. Visual properties of neurons in inferotemporal cortex of the Macaque. *J Neurophysiol.* 1972; 35:96–111. [PubMed: 4621506]
- Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol.* 1962; 160:106–154. [PubMed: 14449617]
- Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. *J Physiol.* 1968; 195:215–243. [PubMed: 4966457]
- Huetz C, Gourevitch B, Edeline JM. Neural codes in the thalamocortical auditory system: from artificial stimuli to communication sounds. *Hear Res.* 2011; 271:147–158. [PubMed: 20116422]
- Hugdahl K, Thomsen T, Ersland L, Rimol LM, Niemi J. The effects of attention on speech perception: an fMRI study. *Brain Lang.* 2003; 85:37–48. [PubMed: 12681347]
- Jongman A. Duration of frication noise required for identification of English fricatives. *J Acoust Soc Am.* 1989; 85:1718–1725. [PubMed: 2708688]
- Kawahara H. Speech representation and transformation using adaptive interpolation of weighted spectrum. *IEEE Trans Acoustics Speech Signal Processing.* 1997; 2:1303–1306.
- Kent, RD.; Read, C. The acoustic analysis of speech. Singular; 2001.
- Kilgard MP. Harnessing Plasticity to Understand Learning and Treat Disease. *Trends in Neuroscience.* 2012; 35:715–722.
- Kilgard MP, Merzenich MM. Cortical map reorganization enabled by nucleus basalis activity. *Science.* 1998; 279:1714–1718. [PubMed: 9497289]
- Klatt DH. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America.* 1980; 67:971–995.
- Kluender KR, Diehl RL, Killeen PR. Japanese-Quail Can Learn Phonetic Categories. *Science.* 1987; 237:1195–1197. [PubMed: 3629235]
- Knudsen EI, Konishi M. Space and frequency are represented separately in auditory midbrain of the owl. *J Neurophysiol.* 1978; 41:870–884. [PubMed: 681991]
- Kuhl PK, Miller JD. Speech-Perception by Chinchilla - Voiced-Voiceless Distinction in Alveolar Plosive Consonants. *Science.* 1975; 190:69–72. [PubMed: 1166301]
- Langner G, Albert M, Briede T. Temporal and spatial coding of periodicity information in the inferior colliculus of awake chinchilla (*Chinchilla laniger*). *Hear Res.* 2002; 168:110–130. [PubMed: 12117514]
- Langner G, Schreiner CE. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J Neurophysiol.* 1988; 60:1799–1822. [PubMed: 3236052]
- Lewicki MS. Intracellular characterization of song-specific neurons in the zebra finch auditory forebrain. *J Neurosci.* 1996; 16:5855–5863. [PubMed: 8795637]
- Lewicki MS, Konishi M. Mechanisms underlying the sensitivity of songbird forebrain neurons to temporal order. *Proc Natl Acad Sci U S A.* 1995; 92:5582–5586. [PubMed: 7777552]
- Lieberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M. Perception of the speech code. *Psychol Rev.* 1967; 74:431–461. [PubMed: 4170865]
- Loizou PC, Dorman M, Tu ZM. On the number of channels needed to understand speech. *Journal of the Acoustical Society of America.* 1999; 106:2097–2103. [PubMed: 10530032]
- Middlebrooks JC, Zook JM. Intrinsic organization of the cat's medial geniculate body identified by projections to binaural response-specific bands in the primary auditory cortex. *J Neurosci.* 1983; 3:203–224. [PubMed: 6185655]
- Miller GA, Nicely PE. An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am.* 1955; 27:338–352.
- Narayan R, Grana G, Sen K. Distinct time scales in cortical discrimination of natural sounds in songbirds. *J Neurophysiol.* 2006; 96:252–258. [PubMed: 16571738]
- Nelken I. Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol.* 2004; 14:474–480. [PubMed: 15321068]
- Palombi PS, Caspary DM. Physiology of the young adult Fischer 344 rat inferior colliculus: responses to contralateral monaural stimuli. *Hear Res.* 1996; 100:41–58. [PubMed: 8922979]
- 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.* 2012

- Peterson G, Barney H. Control methods used in a study of the vowels. *J Acoust Soc Am.* 1952; 24:175–183.
- Porter BA, Rosenthal TR, Ranasinghe KG, Kilgard MP. Discrimination of brief speech sounds is impaired in rats with auditory cortex lesions. *Behav Brain Res.* 2011; 219:68–74. [PubMed: 21167211]
- Ramus F, Hauser MD, Miller C, Morris D, Mehler J. Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science.* 2000; 288:349–351. [PubMed: 10764650]
- Ranasinghe KG, Carraway RS, Borland MS, Moreno NA, Hanacik EA, Miller RS, Kilgard MP. Speech discrimination after early exposure to pulsed-noise or speech. *Hear Res.* 2012a; 289:1–12. [PubMed: 22575207]
- Ranasinghe KG, Vrana WA, Matney CJ, Kilgard MP. Neural mechanisms supporting robust discrimination of spectrally and temporally degraded speech. *J Assoc Res Otolaryngol.* 2012b; 13:527–542. [PubMed: 22549175]
- Read HL, Winer JA, Schreiner CE. Modular organization of intrinsic connections associated with spectral tuning in cat auditory cortex. *Proc Natl Acad Sci U S A.* 2001; 98:8042–8047. [PubMed: 11438747]
- 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. *J Exp Anal Behav.* 2003; 80:205–215. [PubMed: 14674729]
- Rennaker RL, Carey HL, Anderson SE, Sloan AM, Kilgard MP. Anesthesia suppresses nonsynchronous responses to repetitive broadband stimuli. *Neuroscience.* 2007; 145:357–369. [PubMed: 17207583]
- Schreiner CE, Langner G. Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *J Neurophysiol.* 1988; 60:1823–1840. [PubMed: 3236053]
- Schreiner CE, Mendelson JR. Functional topography of cat primary auditory cortex: distribution of integrated excitation. *J Neurophysiol.* 1990; 64:1442–1459. [PubMed: 2283536]
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech Recognition with Primarily Temporal Cues. *Science.* 1995; 270:303–304. [PubMed: 7569981]
- 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; 34:1823–1838. [PubMed: 22098331]
- Sinex DG, Chen GD. Neural responses to the onset of voicing are unrelated to other measures of temporal resolution. *J Acoust Soc Am.* 2000; 107:486–495. [PubMed: 10641656]
- Steinschneider M, Arezzo J, Vaughan HG Jr. Speech evoked activity in the auditory radiations and cortex of the awake monkey. *Brain Res.* 1982; 252:353–365. [PubMed: 7150957]
- Steinschneider M, Fishman YI, Arezzo JC. Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. *J Acoust Soc Am.* 2003; 114:307–321. [PubMed: 12880043]
- Steinschneider M, Volkov IO, Fishman YI, Oya H, Arezzo JC, Howard MA 3rd. Intracortical responses in human and monkey primary auditory cortex support a temporal processing mechanism for encoding of the voice onset time phonetic parameter. *Cereb Cortex.* 2005; 15:170–186. [PubMed: 15238437]
- Steinschneider M, Volkov IO, Noh MD, Garell PC, Howard MA 3rd. Temporal encoding of the voice onset time phonetic parameter by field potentials recorded directly from human auditory cortex. *J Neurophysiol.* 1999; 82:2346–2357. [PubMed: 10561410]
- Suga N. Principles of auditory information-processing derived from neuroethology. *J Exp Biol.* 1989; 146:277–286. [PubMed: 2689566]
- Turner JG, Hughes LF, Caspary DM. Divergent response properties of layer-V neurons in rat primary auditory cortex. *Hear Res.* 2005; 202:129–140. [PubMed: 15811705]
- Watanabe T. Responses of the cat's collicular auditory neuron to human speech. *J Acoust Soc Am.* 1978; 64:333–337. [PubMed: 711999]
- Wong SW, Schreiner CE. Representation of CV-sounds in cat primary auditory cortex: intensity dependence. *Speech Communication.* 2003; 41:93–106.

Young ED. Neural representation of spectral and temporal information in speech. *Philos Trans R Soc Lond B Biol Sci.* 2008; 363:923–945. [PubMed: 17827107]

1. Primary auditory cortex (A1) evokes more diverse firing to speech than inferior colliculus (IC).
2. IC neurons are primarily driven by the spectral energy of speech sounds.
3. A1 neurons respond to multiple acoustic properties of speech sounds

Temporal Levels (Low pass filter cutoff)	4 Hz (16 bands)			
	16 Hz (16 bands)			
	64 Hz (16 bands)			
	16 bands 256 Hz	8 bands (256 Hz)	4 bands (256 Hz)	2 bands (256 Hz)
Undegraded	Spectral Levels (Number of Bands)			

Figure 1. Levels of Degradation

A noise vocoder was used to degrade speech spectrally and temporally. Four spectral levels (along horizontal axis) were defined by the number of band pass filters used in the noise vocoder as 16, 8, 4, and 2. Increased number of bands provided increased spectral resolution. Four temporal levels (along the vertical axis) were defined by the low pass filter cutoff frequency as 256 Hz, 64 Hz, 16 Hz and 4 Hz. Higher filter cut off frequency provided higher temporal resolution. The four spectrally degraded levels maintained the temporal resolution at 256 Hz, and the four temporally degraded levels maintained the spectral resolution at 16 bands.

Figure 2. Speech evoked activity in IC and A1

(a) Spectrograms of octave shifted, undegraded speech syllables of 'Dood', 'Dad' and 'Tad'. The intensity of the gray scale is proportional to the intensity of the signal. (b) Population PSTH (average of 118 sites) showing the IC response to each of the speech syllables whose spectrograms are shown in (a) above. IC neurons (multi-units) generated a sharp onset response for the initial consonant and a strong sustained response for the relatively steady state vowel component of the syllable. (c) Population PSTH (average of 120 sites) showing the A1 response to each of the speech syllables whose spectrograms are shown in (a) above. A1 neurons (multi-units) generated a sharper onset response for the initial consonant and a relatively low sustained response compared to IC, for the steady state vowel component of the syllable. Spike rate of IC was higher than A1. IC=inferior colliculus; A1=primary auditory cortex.

Figure 3. Neurons with similar frequency tuning generated more diverse spike count profiles in A1 compared to IC

Spike counts of the first 300 ms window generated by two IC sites tuned to 1 kHz for the four vowel stimuli ('dad', 'dud', 'deed', 'dood') (a) and two A1 sites tuned to 1 kHz (b). IC and A1 the spike rate was normalized to the average firing (300 ms) of each population. The two IC sites tuned to the same characteristic frequency (1 kHz) generated highly correlated response (c), whereas the response patterns of the two A1 sites tuned to the same characteristic frequency (1 kHz) were not correlated (d). Another example of spike count profiles generated in response to the four vowel stimuli by two IC sites tuned to 10 kHz (e) and two A1 sites tuned to 10 kHz (f). The two IC sites tuned to the same characteristic frequency (10 kHz) generated highly correlated responses (g), whereas the response patterns of the two A1 sites tuned to the same characteristic frequency (10 kHz) were not correlated (h). (i) Histogram demonstrating the distribution correlations between all IC pairs ($n=117$) where the two sites are tuned to the same characteristic frequency. More than 60% of pairs (71 out of 117) generated correlation coefficients above 0.70. (j) Histogram distribution of correlations between all A1 pairs ($n=118$) where the two sites are tuned to the same

characteristic frequency. Less than 20% of pairs (23 out of 119) generated correlation coefficients above 0.70. Corr coefficient=correlation coefficient. IC=inferior colliculus; A1=primary auditory cortex

Figure 4. Neurons with similar frequency tuning generated more diverse spike timing patterns in A1 compared to IC

Spike timing pattern of the first 40 ms onset window generated by two IC sites tuned to 1 kHz for the four consonant stimuli ('d', 'b', 's', 't') (a) and two A1 sites tuned to 1 kHz (d). The two IC sites tuned to the same characteristic frequency (1 kHz) generated more correlated responses compared to the two A1 sites tuned to the same characteristic frequency (1 kHz). Another example of spike timing profiles generated in response to the four consonant stimuli by two IC sites tuned to 10 kHz (b) and two A1 sites tuned to 10 kHz (e). The two IC sites tuned to the 10 kHz generated more correlated spike timing patterns, compared to the two A1 sites tuned to 10 kHz. (c) Histogram demonstrating the distribution of spike timing pattern correlations between all IC pairs where the two sites are tuned to the same characteristic frequency. More than 60% of pairs (73 out of 117) generated correlation coefficients above 0.5. (f) Distribution of correlations of spike timing patterns between all A1 pairs where the two sites are tuned to the same characteristic frequency. Less than 50% of pairs (56 out of 119) generated correlation coefficients above 0.5. Corr coefficient=correlation coefficient. IC=inferior colliculus; A1=primary auditory cortex

Figure 5. Spectral degradation changed neural distinctiveness of vowels equally in IC and A1
 (a) The spike count profiles (spatial patterns), plotting the number of spikes fired during the first 300 ms of speech evoked response by IC neurons (left panels) and A1 neurons (right panels). X-axis is arranged according to the characteristic frequency of each neuron. Y axis plots the total number of spikes (averaged across 20 repeats) for the first 300 ms from the voicing onset of each recording site. For demonstration purposes number of spikes was smoothed across 10 adjacent sites. The example shows the spike patterns generated for two of the four vowels used ('dad' and 'dood'). Responses for the two undegraded vowels and for the four spectrally degraded stages are shown from top to bottom. The degree of change in spatial patterns as the stimuli was degraded is similar in IC and A1. Euclidean Distance between spatial patterns of the 6 vowel pairs ('dad'-'dood', 'dad'-'dud', 'dad'-'deed', 'dood'-'dud', 'dood'-'deed' 'dud'-'deed') was measured at each spectral level. (b) The percentage of change in Euclidean distance at each increasing step in degradation. 16b/

Undeg=ED at 16 bands as percent of Euclidean Distance at undegraded; 8b/16b=ED at 8 bands as percent of Euclidean Distance at 16 bands; 4b/8b=ED at 4 bands as percent of Euclidean Distance at 8 bands; 2b/4b=ED at 2 bands as percent of Euclidean Distance at 4 bands; ED=Euclidean distance; IC=inferior colliculus; A1=primary auditory cortex

Figure 6. Temporal degradation changed neural distinctiveness of consonants equally in IC and A1

(a) Neurograms depicting the evoked activity of the first 40 ms onset response by IC neurons (left panels) and A1 neurons (right panels). The example shows the spatiotemporal patterns generated for two of the four consonants used ('d' and 't'). Neurograms of the two undegraded consonants and those for the four temporally degraded stages are shown from top to bottom. The degree of change in spatiotemporal activity patterns as the stimuli are degraded is similar in IC and A1. Euclidean Distance between neurograms of the 6 consonant pairs ('d'-b', 'd'-s', 'd'-t', 'b'-t', 'b'-s', 't'-s') was measured at each temporal level. (b) The percentage of change in Euclidean distance at each increasing step in degradation. 256 Hz/Undeg=ED at 256 Hz as percent of Euclidean Distance at undegraded;

64 Hz/256 Hz=ED at 64 Hz as percent of Euclidean Distance at 256 Hz; 16 Hz/64 Hz=ED at 16 Hz as percent of Euclidean Distance at 64 Hz; 4 Hz/16 Hz=ED at 4 Hz as percent of Euclidean Distance at 16 Hz; ED=Euclidean distance; IC=inferior colliculus; A1=primary auditory cortex

Figure 7. Spectral and temporal degradation of vowels affect A1 diversity and IC diversity differently

Correlation coefficient depicts the average correlation of the vowel spike counts between neurons clustered into a half octave from their characteristic frequency (open markers) and neurons that are tuned to a frequency half octave away from their characteristic frequency (shaded markers), at each spectral level (a) and temporal level (b). To the left of the dotted line are the noise vocoded (degraded) stimuli. Response strength was estimated by the number of spikes generated by IC neurons and A1 neurons in response to vowels (300 ms window), at each spectral (c) level and temporal level (d). The spike rate of IC and A1 was normalized to the average firing rate (300 ms window) of each population (i.e. all IC and A1 responses, respectively). IC=inferior colliculus; A1=primary auditory cortex; Undeg=Undegraded

Figure 8. Tuning properties of neurons contribute to diversity of A1 neural activity patterns evoked by vowels

Correlation coefficient depicts the average correlation of the vowel spike counts between neurons clustered into a half octave from their characteristic frequency in IC (a),(c) and A1(b),(d). Open markers in (a) and (b) indicate the neurons that are within 1 octave of each other in their bandwidth at 10 dB threshold level (BW10). Shaded markers in (a) and (b) indicate neurons that are beyond 1 octave of each other in BW10. Open markers in (c) and (d) indicate the neurons that are within 1 ms of each other in their latency response and the shaded markers in (c) and (d) indicate neurons that are beyond 1 ms from each other in their latency response. BW10=bandwidth at 10dB threshold; Latency diff= latency difference; IC=inferior colliculus; A1=primary auditory cortex; Un=Undegraded

Figure 9. Spectral degradation has little effect on neural diversity of IC and A1 while temporal degradation produces significant changes

Correlation coefficient depicts the average correlation of the consonant spike timing activity between neurons clustered into a half octave from their characteristic frequency (open markers) and neurons that are tuned to a frequency half octave away from their characteristic frequency (shaded markers), at each spectral level (a) and temporal level (b). To the left of the dotted line are the noise vocoded (degraded) stimuli. Response strength was estimated by the number of spikes generated by IC neurons and A1 neurons in response to consonants (40 ms window), at each spectral (c) level and temporal level (d). IC and A1 the spike rate was normalized to the average firing (40 ms onset window) of each population. IC=inferior colliculus; A1=primary auditory cortex; Undeg=Undegraded

Figure 10. Tuning properties of neurons contribute similarly to A1 and IC diversity of neural activity patterns evoked by consonants

Correlation coefficient depicts the average correlation of the spike timing activity evoked by consonants between neurons clustered into a half octave from their characteristic frequency in IC (a),(c) and A1(b),(d). Open markers in (a) and (b) indicate the neurons that are within 1 octave of each other in their bandwidth at 10 dB threshold level (BW10). Shaded markers in (a) and (b) indicate neurons that are beyond 1 octave of each other in BW10. Open markers in (c) and (d) indicate the neurons that are within 1 ms of each other in their latency response and the shaded markers in (c) and (d) indicate neurons that are beyond 1 ms from each other in their latency response. BW10=bandwidth at 10dB threshold; Latency diff=latency difference; IC=inferior colliculus; A1=primary auditory cortex; Un=Undegraded

Table 1

Tuning Curve Characteristics of IC and A1 neural populations

	IC	A1
Number of multi-units		
Total	118	120
CF < 6 kHz	36	58
CF > 6 < 11 kHz	34	39
CF > 11 kHz	48	23
Bandwidth		
BW 10	1.83 (0.12)	1.60 (0.07)
BW 20	2.32 (0.12)	2.30 (0.08)
BW 30	2.69 (0.13)	2.96 (0.10)
BW 40	3.12 (0.13)	3.61 (0.10)
Average Firing Rate	4.01 (0.19)	2.41 (0.08)
Latency	14.10 (0.21)	16.77 (0.08)

CF=characteristic frequency; BW=bandwidth; numbers within parenthesis show standard errors.