Despite its ubiquity in music and speech, the cortical origins and neural mechanisms underlying pitch perception are still not well understood. While some researchers have shown cortical correlates of pitch in lateral Heschl’s gyrus (HG) (Patterson et al., 2002; Penagos et al., 2005; Puschmann et al., 2010), others argue for a locus in planum temporale (PT) (Hall and Plack, 2009). A more distributed representation of pitch along the HG has also been argued (Griffiths et al, 2010), where different parts of HG have specific functions within a pitch system (Kumar et al., 2011). Similarly, evidence from animal studies is also not consistent in that some studies show a single locus (Bendor and Wang, 2005), while others (Bizley et al., 2009) suggest a distributed representation of pitch.

Previous fMRI studies have generally used sustained sounds of several seconds duration within a sparse temporal sampling acquisition design. This design assesses only the steady-state response to pitch and lacks information about the temporal dynamics of the response. Here, we make use of an adaptation fMRI paradigm (Grill-Spector and Malach, 2001) to elucidate areas in human auditory cortex that show repetition suppression (RS) to repeated presentations of a pitch stimulus. Importantly, we seek areas that show stronger RS for pitch than spectrally-matched control sounds.

In separate experiments, two types of pitch evoking stimuli (harmonic complex sounds and regular click trains, 1.8 s) and the corresponding spectrally matched noise were presented four times in a trial. The decay of activity with the number of repetitions was modelled using an exponential decay function $\exp(k \cdot n)$, where $k$ is the decay constant and $n$ is the number of repetitions. We determined the value of the decay constant $k$ for each voxel of the auditory cortex by using a Taylor series expansion around a nominal value of $k$. Initial data show responses to pitch that decay faster than the response to noise in mid to lateral HG, extending posterior-laterally into Heschl’s sulcus.

The data provide further support for specific pitch mechanisms in mid to lateral HG.

References

Puschmann et al. (2010). Neuroimage, 2010, 49(2), 1641-9
Prototype-based cerebral representation of voice identity

Pascal Belin & Marianne Latinus

Voice Neurocognition Laboratory, Institute of Neuroscience and Psychology, University of Glasgow

How different voices are represented in the brain remains largely mysterious. We provide evidence of prototype-based representation of voice identity, using both behavioural identification tasks and fMRI measures of auditory cortex activity. In experiments using voice identity aftereffects (Zaske et al), we show that "anti-voice" adaptors - stimuli opposite to an original voice relative to the average voice, and perceived as a different identity - induce larger aftereffects than other adaptors, consistent with a special role of the average voice in representing voice identity. By representing individual voices in a 3D acoustical space reflecting the main acoustical dimensions of voices - f0, formant frequency, harmonicity - we show, for both male and female voices, that the Euclidean distance between a voice and the average of all voices ("distance to mean") correlates with that voice's perceived distinctiveness. Moreover, voices' distance to mean explain a significant portion of the variance in fMRI signal of temporal voice areas (TVA) response (up to 25% of the variance, r=0.5), with voices farther away from the average eliciting greater TVA activity. These results provide converging evidence that voices are represented in the brain as a function of their distance in a multidimensional "voice space" to an internal voice prototype well approximated by the morphed average of a large number of voices: a voice with average f0 and formant frequencies, but with an unnaturally high harmonicity.
Biasing the content of hippocampal replay during sleep using task-related sounds

Daniel Bendor and Matt Wilson

Picower Institute for Learning and Memory
MIT
Cambridge, MA 02139

Cortico-hippocampal interactions during non-REM sleep play an important role in memory consolidation. The hippocampus acts as a temporary storage of information related to our autobiographical experiences, and during sleep this information is transferred to neocortex for long-term storage. Several recent studies in humans have demonstrated an improvement in memory consolidation when task-related cues (acoustic or olfactory) were presented during a non-REM sleep session following training [Rasch et al. 2007, Rudoy et al. 2009]. One possible neural mechanism by which the consolidation of specific memories could be enhanced is replay. During replay, neural ensembles in the hippocampus and cortex encoding a previous experience are reactivated together. Coordinated cortico-hippocampal replay provides a possible neural mechanism by which a memory stored in the hippocampus can be transferred to cortex. According to this model, the strength of memory consolidation for a particular experience will scale with the degree to which that encoded experience replays.

We examined in rodents whether memory improvements in these tasks were due to a bias in the content of hippocampal replay following the presentation of task-related cues. Rats were trained on an auditory-spatial association task, and during a post-training sleep session, task-related sounds were played to the rat while recording from ensembles of neurons in the hippocampus (dorsal CA1). We observed that the content of replay activity occurring after the presentation of a task-related sound was biased towards the spatial locations associated with that sound during the behavioral task. These data suggest that biasing replay content towards a particular experience causes that experience to be more strongly consolidated.
Title: Psychophysical correlates of co-modulation masking release in a macaque
Authors: Sharath Bennur & Yale E Cohen

Abstract: Perception of communication signals in noisy environments is a critical function of the auditory system. Our ability to perceive a communication signal in a noisy environment is aided by certain properties of the background noise: hearing thresholds in co-modulated noise are lower than in unmodulated noise. This effect is called “co-modulation masking release”. It is hypothesized that the fluctuating nature of the co-modulated background gives the auditory system access to “glimpses” of the stimulus, therefore enabling lowered detection thresholds. While this effect, along with other forms of masking and release from masking, has been studied extensively in humans, to date, there have not been any studies using non-human primates.

Here, we describe a novel tone-in-noise detection task. Monkeys were trained to report the presence of a tone embedded in either co-modulated or unmodulated noise. On a trial-by-trial basis, the audibility of the tone was varied systematically by changing the intensity of the tone, relative to a background noise with a fixed intensity. When the monkey heard the tone, he released a lever to obtain a juice reward. Behavioral performance was quantified with a psychophysical curve and a d’ analysis. We report how performance varied as a function of tone frequency, tone intensity, and noise type (co-modulated or unmodulated). These results provide a novel behavioral task that can be used to study the link between hearing in noisy environments and its neural underpinnings.
Primary sensory cortex predicts the utility of specific sensory information in a behavioral learning set by enhancing the cortical representation of the critical signal.

K. M. BIESZCZAD*, & N. M. WEINBERGER

Cntr Neuro. of Learn. & Memory and Dept. of Neuro. & Behav., Univ. California Irvine, Irvine, CA.

The representation of acoustic frequency in the tonotopic map of primary auditory cortex (A1) is a convenient model for “item-specific” information storage in the cerebral cortex. For example, tonotopy in A1 encodes the behavioral relevance of a signal frequency by enhancing its cortical representation; the importance of the acoustic-frequency detail of an experience can be represented in the number of A1 cells that become closely tuned to that frequency (Rutkowski & Weinberger, 2005) which may increase its strength memory (Bieszczad & Weinberger, 2010). Here we investigated how A1 might represent the importance of more than one frequency that have become signals in a learning set. We examined the dynamics of receptive field plasticity as rats implanted with an array of recording electrodes learned to bar-press for rewards to one tone-signal (5 kHz) and in a second phase, to discriminate between this and another spectrally-distant signal-frequency (5 vs. 12 kHz). Multiple recordings from A1 were made prior to daily training sessions to characterize frequency response areas (FRAs) at each site. FRAs were compared to daily performance to investigate neural predictive correlates of behavioral frequency-discriminations. We report the development of A1 receptive field plasticity in a behavioral challenge, “partial reversal discrimination”, when an original single CS+ (5 kHz), was made the unrewarded CS- with a new CS+ (12 kHz) introduced for discrimination. A frequency-discrimination learning set was established by subsequent consecutive reversals in which the signal-frequency that was rewarded (CS+) or unrewarded (CS-) was reversed. The findings reveal that current reward valence and prior experience combine to reflect high-order instantiation of frequency-specific A1 plasticity that serves the function of memory for useful auditory signals.
Auditory recognition memory correlates and stimulus-selectivity of local field potentials in primate lateral prefrontal cortex

James Bigelow\textsuperscript{1}, Bethany Plakke\textsuperscript{2}, and Amy Poremba\textsuperscript{1,3}, \textsuperscript{1}Department of Psychology, Division of Behavioral and Cognitive Neuroscience, University of Iowa, Iowa City, IA 52242, \textsuperscript{2}University of Rochester, Rochester, NY, 14642, \textsuperscript{3}Neuroscience Program, University of Iowa, Iowa City, IA 52242.

The lateral prefrontal cortex (PFC) is central to the neural circuitry underlying integration of sensory events and internal states guiding prospective goal-directed actions involving short-term retention of information (Miller & Cohen, 2001). Although the lateral PFC has prominent visual, auditory, and somatosensory domains (Barbas et al., 2002), the vast majority of experiments examining PFC mnemonic processes use visual stimuli. Herein, we present findings from local field potential (LFP) recordings in the lateral PFC of two rhesus monkeys performing a delayed matching-to-sample (DMTS) task. Auditory stimuli were selected from a collection of diverse auditory stimuli. Twelve different stimulus sets were created with one exemplar from each of the following eight categories: conspecific monkey vocalizations, human vocalizations, animal vocalizations, natural environmental sounds, pure tones, synthetic sounds, music clips, and band-passed white noise. Each experimental session began with a passive-listening phase wherein one of the stimulus sets was presented. The same stimuli were then presented for discrimination during the DMTS task. Each trial consisted of a sample stimulus, followed by a 5s retention delay, followed by an identical (match trials) or different (non-match trials) test stimulus. Correct responses were defined as a button press following matching stimuli, and withholding button presses after non-matching stimuli. Following the task, a larger passive listening block was presented in which all stimuli from each of the twelve stimulus sets. We recorded LFP at 237 sites in Area 46, of which 99 were responsive to auditory stimuli and further analyzed. Many sites exhibited stimulus selectivity in the form of differential evoked responses between stimuli. Across recording sites and stimuli, greater evoked responses were observed during DMTS than passive listening, indicating that task-relevant auditory stimuli increase neuronal activation and/or synchrony in the PFC. Further, for many sites, differential responses were elicited by matching vs. non-matching test stimuli corroborating our “match enhancement” patterns observed in PFC spiking activity during the same auditory recognition memory encoding (Plakke et al., 2010). Some changes were apparent in specific frequency bands at a limited number of sites, e.g., during the memory delay interval, gamma band activity decreasing as the interval progressed. These results suggest that the lateral PFC is engaged in processing and retaining auditory cues that are needed to accomplish a prospective task perhaps providing a common processing mechanism across multiple modalities.

Support: Startup funds from the Univ. of Iowa and NIDCD grant DC0007156 to A.P.

Key words: cortex, rhesus macaque, monkey, working memory, sound, gamma, area 46
Neural correlates of pitch discrimination during passive and active listening


Abstract: We have obtained electrophysiological recordings from the auditory cortex of freely moving ferrets while they perform a pitch discrimination task. Neural data were obtained using microelectrode arrays (Neuralynx, WARP-16 devices), which were chronically implanted into the auditory cortex of ferrets, following methods adapted from those used by Eliades and Wang (2008). Neural signals were recorded from animals while they were tested, in twice-daily sessions, over the course of a year. Over this time period the electrodes were independently advanced through the auditory cortex approximately weekly. Animals were trained in a two-alternative forced choice task that required them to judge whether the second of two artificial vowel sounds (the “target”) was higher or lower in pitch than a preceding “reference” sound (Walker et al 2009). Local field potential (LFP) and spike data were recorded, both during the behavioral task, and immediately before or after, when the animal was passively listening to the same sounds. ROC analysis revealed that high and low pitch targets often differentially modulated the LFP power observed during the target sound. The LFP power was also often predictive of the animal’s behavioral response. In many cases the LFP power discriminated the animal’s judgment with a slightly greater accuracy than it did the pitch of the acoustic stimulus. We examined the timescale over which the LFP power became informative about either the pitch of the sound or the animal’s behavioral response and observed that choice probabilities emerged later in the neural response than stimulus sensitivity. LFPs recorded in the same animals while they were either awake and passively listening, or asleep, differed considerably in their amplitude, shape and tuning from those recorded when the same animals were actively discriminating the target pitch. ROC analysis applied to responses collected in the passively listening animal frequently revealed these responses to be less informative about the target pitch than those recorded during behavior.
Task demands and motivation affect neuronal activity in the auditory cortex of nonhuman primates

Michael Brosch, Mikhail Babanin, Elena Selezneva, Ying Huang, Henning Scheich

Leibniz Institut für Neurobiologie
Brennekestrasse 6
39118 Magdeburg, Germany

More recent findings have demonstrated that auditory cortex also subserves functions other than sound perception and auditory memory. Here we addressed the question how procedural and motivational aspects of an audio-motor task affect neuronal firing in auditory cortex. Two long-tail macaques were trained to perform a simple detection task in which subjects had to report, by bar release, the termination of an audiovisual stimulus. The procedural aspect was singled out by requiring the subjects to attend either to the auditory modality or to the visual modality. The motivational aspect was singled out by having the subjects perform the two tasks with different sizes of rewards. Analysis of multiunit activity from several hundred recording sites in primary and secondary auditory fields revealed that both aspects were most strongly reflected in slow changes (in the order of seconds) of neuronal activity in auditory cortex. The firing changes typically started after onset of the audiovisual stimulus and continued throughout the entire duration of this stimulus. These changes were seen both when the monkeys attended to the auditory and to the visual modality, but were generally stronger and seen in more units in the former condition. Reward expectation was also reflected in the slope of the changes: typically, expectation of a large reward resulted in stronger firing changes than a small reward. Our results support and extend previous accounts of cognitive aspects of auditory cortex function.
Although already introduced more than 120 years ago, the Golgi Silver-impregnation method still provides detailed information about the morphology of neurons and about the architecture of brain structures that is only hardly available by means of other methods. Here, we investigated the anatomy of the subcortical auditory pathway in the Mongolian gerbil (*Meriones unguiculatus*), a common animal model in auditory research, using the Golgi-Cox Method.

Within the medial geniculate body, a ventral (MGv), dorsal (MGd) and medial (MGm) division can be identified. The MGv consists of a laminated lateral part (*pars lateralis*), a circular medial part (*pars ovoidea*) and a rostral pole. The most conspicuous neurons are medium to large-sized tufted cells. Within the MGd, a deep dorsal subdivision (DD) and the suprageniculate nucleus (SG) are apparent. The DD contains densely packed tufted neurons with bushy dendritic arbors and round cells with radiate dendrites. In the SG, large radiate or stellate neurons can be frequently found. The MGm harbors a diverse population of densely packed neurons, including “magnocellular” and medium to large-sized tufted fusiform cells and small cells with tufted or radiate dendrites.

The inferior colliculus of the gerbil can be parcellated into a central nucleus (CIC), a dorsal (DCIC) and an external cortex (ECIC). Within the laminated CIC, we found disk-shaped and stellate cells. The DCIC and ECIC display a cortex-like architecture with three layers. Generally, there is a trend of an increasing cell size and packing density of the mostly bipolar, multipolar or triangular neurons towards deeper layers within both divisions.
The dorsal nucleus of the lateral lemniscus contains a variety of round, multipolar and elongated cells. The intermediate nucleus contains diverse multipolar and horizontal cells and the bipartite ventral nucleus contains medium-sized to very large cells with multipolar or vertically elongated perikarya and small globular cells.

The most prominent divisions of the superior olivary complex are the lateral (LSO) and medial superior olive, the medial nucleus of the trapezoid body and the superior periolivary nucleus. The LSO, for example, harbors medium-sized to large bipolar to fusiform neurons with disk-shaped dendritic fields and multipolar neurons with more radiate dendritic fields.

The gerbil's cochlear nuclear complex consist of a 4-layered dorsal nucleus, containing, e.g., cartwheel, bipolar, pyramidal, granule and giant neurons, as well as of an anteroventral and of a posteroventral nucleus, both containing spherical and globular bushy, octopus, multipolar and small cells.
Arc in sensory cortex is necessary for learning and physiological plasticity.

E. CARPENTER-HYLAND, A. VAZDARJANOVA, D. BLAKE;
Brain and Behavior Discovery Inst., Med. Col. of GA, AUGUSTA, GA

Enhancing neural representations of learned stimuli in primary sensory cortex is critical for perceptual learning. The molecular signals and neural codes that support perceptual learning were the focus of study in rats being trained to perform a tone detection task. Arc mRNA translation was blocked locally in primary auditory cortex (A1) during the third day of acoustic detection learning by infusing an anti-sense or scrambled oligonucleotide. While Arc block had no effect on behavioral performance the day of infusion, the following day’s performance was as though the third session was not consolidated and learning from the first two sessions was erased. In contrast, Arc block 10 hours following the third session had no effect on behavioral performance the following day. Reconstructing electrophysiological activity in A1 showed that 23 hours following Arc block average response to target were increased, but responses occurred in under half as much cortical area. Detection models that include either rate or coincidence coding explain how increases in the cortical area of a response can improve its perceptual salience.
Despite numerous advances in understanding the structural, connectional, and functional characteristics of auditory cortex, a consistent model of information processing linking laminar connectivity and functional properties has not yet been established. While work in visual and somatosensory cortices of various species has demonstrated a pattern of activation defined by granular→supragranular→infragranular neuronal engagement, investigations in the auditory cortex of the mouse, rat, Mongolian gerbil, and guinea pig have provided evidence of a serial activation pattern moving from infragranular to supragranular layers. In this investigation we quantified single-unit neuronal responses across cortical laminae during exposure to pure tones, noise bursts, FM sweeps, and conspecific vocalizations. Neuronal activity measured simultaneously across A1 laminae in the adult domestic cat revealed a sequential pattern of neuronal activation moving from deep to superficial layers of auditory cortex. Furthermore, response time cross-correlation analysis identified a directional flow of activation (granular→supragranular→infragranular) across layers. Our results suggest that acoustic information in A1 induces neuronal responses in deep laminae prior to neuronal activation in superficial laminar regions. These results are comparable to patterns of activation in auditory cortex identified in rodents, but are in contrast to laminar patterns of activation shown in primary visual and somatosensory cortices.

This work was supported by grants from the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada, and the Canada Foundation for Innovation.
Abstract: Rats produce complex vocalizations in communicating with each other. Over 14 types of distinct calls can be distinguished in their repertoire. However, how these complex vocalizations are encoded in the auditory pathway is not well understood. To learn how the auditory cortex encoded information about rat vocalizations, we presented a library of modified vocalizations to awake rodents and recorded neural activity in the auditory cortex.

To generate a library of acoustic stimuli, we first used the matching pursuit algorithm with a gammatone basis to create a sparse representation of the structure of vocalizations. We next applied statistical analysis on the spectro-temporal relation between the coefficients of the gammatones to create a low-parameter description of the cross-channel structure of the vocalization. We determined the statistical variables that accounted for the diversity in the vocalizations and constructed a library of artificial stimuli, in which the specific parameters were modulated.

Neurons in the auditory cortex responded reliably to the distinct vocalizations. We assayed how well a generalized linear non-linear model (GLNM) fitted from a random chirp stimulus predicted their responses. Using the responses to modified vocalizations, we extended the GLNM to incorporate a term that explicitly encodes the statistical correlations across the filter bank to improve model predictions to vocalizations.
Reference frame of visual and auditory signals in the primate frontal eye fields

V. CARUSO¹, D. PAGES², J. M. GROH³;
¹Duke University, Ctr. For Cognitive Neurosci., Durham, NC; ²Duke University, Ctr. for Cognitive Neurosci., Durham, NC; ³Duke University, Ctr. for Cognitive Neuroscience, Dept. of Psychology and Neuroscience, Dept. of Neurobiology, Duke Institute for Brain Science, Durham, NC

Abstract: We localize events in the world by integrating signals from multiple sensory modalities. Visual and auditory stimuli are initially detected in different reference frames: the retina provides eye-centered information about stimulus location whereas the ears provide head-centered information about stimulus location via binaural and spectral cues. The Frontal Eye Fields, one of the cortical areas involved in the voluntary control of saccades, is known to respond to visual and auditory stimuli. Previous work (Russo and Bruce, J Neurophys. 1994) suggested that auditory signals in the FEF depend on sound location with respect to the eyes. Whether this representation is eye-centered (receptive fields move in lock step with the eyes) or in some hybrid reference frame intermediate between head- and eye-centered coordinates has not been determined.

We recorded the activity of single cells in the FEF of a rhesus monkey making saccades from different initial fixation positions to visual or auditory targets. Preliminary results from 62 recordings sites show that the auditory reference frame just after target onset ranges from head-centered to hybrid (mixed eye- and head-centered) coordinates. By the time the movement is generated, the representation has shifted towards eye-centered coordinates, and has become more similar to the visual representation in this structure. Overall, the results in the FEF were similar to our recent findings in the SC (Lee and Groh, SFN 2009), but different from area LIP, in which the hybrid reference frame persists for both modalities through the sensory and motor intervals. (Mullette-Gillman et al., J Neurophys 2005, Cerebral Cortex 2009).
Differential sensitivity to appearing and disappearing objects in complex acoustic scenes

Francisco Cervantes-Constantino\textsuperscript{a}, Makio Kashino\textsuperscript{b} and Maria Chait\textsuperscript{a}

\textsuperscript{a} UCL Ear Institute, 332 Gray’s Inn Rd, London, WC1X 8EE, UK
\textsuperscript{b} NTT Communication Science Laboratories, NTT Corporation, Atsugi, Japan

The ability to detect and quickly respond to changes in our surroundings is crucial for survival. Hearing is hypothesized to play a pivotal role in this process by serving as an ‘early warning device’, rapidly directing attention to unexpected events in the environment. Here we report on a series of psychophysical experiments in which we investigated listeners’ sensitivity to various changes in complex acoustic scenes. We created artificial ‘scenes’ populated by multiple pure-tone components (4-14 components). To ensure they are perceived as different/separate items, they were also characterized by a unique square wave amplitude modulation rate, which mimics temporal properties characteristic of many natural sounds. Importantly, these scenes lacked semantic attributes, which may have been a limiting factor in interpreting previous change detection studies, thus allowing us to probe low level, pre-semantic, processes involved in auditory change perception.

Our results reveal a fundamental difference between ‘appear’ and ‘disappear’ events. Listeners are exceptionally sensitive to change events manifested as object appearance: change detection and identification are at ceiling, response times are short, with little effect of scene-size. In contrast, listeners have difficulty detecting, and still more so identifying, disappearing objects, even in small, acoustically simple, scenes: Performance rapidly deteriorates with growing scene-size, response times are slow, and even when change is detected, the changed component is rarely successfully identified. We also measured change detection performance when a 200 ms noise-gap was inserted at the transition. The data show that gaps adversely affected the processing of item appearance but not disappearance. The qualitatively different performance patterns, which are not entirely consistent with explanations in terms of adaptation or differences in on- and off- tuned cell populations, suggest ‘appear’ and ‘disappear’ change detection may be supported by different mechanisms.
Dynamic faces speed up vocal processing in the auditory cortex of behaving monkeys

Chandramouli Chandrasekaran1,2, Luis Lemus1,2, Asif A Ghazanfar1-3

1Neuroscience Institute, Departments of 2Psychology and 3Ecology & Evolutionary Biology
Princeton University, Princeton, NJ, 08540, USA

We investigated the dynamics of audiovisual integration in auditory cortex of monkeys detecting “coo” vocalizations. Macaque monkeys detected visual, auditory or audiovisual coo vocalizations by monkey avatars in a background of noise (~60 dB) as fast and as accurately as possible. The loudness of the auditory vocalization (i.e. signal to noise ratio or SNR) relative to the noise background was varied; the degree of mouth opening of the avatar was varied to mimic these changes in SNRs. Decreasing SNR of the auditory-only vocalizations led to an increase in reaction times (RTs), whereas decreases in mouth opening size led to a weak increase in RT. Audiovisual RTs were faster than RTs to both auditory- and visual-only conditions. Under these behavioral conditions, we recorded spiking activity and local field potential (LFPs) from ~ 240 cortical sites in the core and lateral belt auditory cortex. We found that, at least in part, that activity in auditory cortex co-varied with behavior. First, a decrease in SNR of the auditory-only vocalization increased response latency and decreased the magnitude of spiking responses. Spiking responses were however absent to visual-only vocalizations. Surprisingly, we found that spiking responses were faster for audiovisual compared to auditory vocalizations—a parallel with decreasing reaction times. Audiovisual spiking benefits without visual spiking activity suggests sub-threshold effects of visual input. Thus, we analyzed the LFP responses to identify the network dynamics that could mediate this speed-up. In keeping with the profiles of spiking activity, LFP responses to audiovisual vocalizations were also faster than to auditory-only vocalizations. In addition, evoked and induced power, as well as inter-trial phase coherence in the 10 – 30 Hz band of the LFP was suppressed for audiovisual compared to auditory-only vocalizations. Suppression was maximal for the largest SNR and decreased with the decrease in SNR. This suggests that visual input into auditory cortex changes the state of network dynamics in this frequency band. We also found that the onset of visual mouth motion led to an increase in the inter-trial phase coherence in the 10 – 30 Hz band immediately after visual onset. These results suggest that during the detection of audiovisual vocalizations, visual cues speed up and alter the dynamics of the circuits in auditory cortex processing vocalizations.
Stimulation of the Amygdala Facilitates Cortical Memory Traces

by Inducing Dual Forms of Representation Plasticity

Chavez, C.M., McGuagh, J.L. & Weinberger, N.M.

Center for the Neurobiology of Learning and Memory,

Department of Neurobiology and Behavior, University of California, Irvine

Activation of the amygdala enhances memory consolidation processes through its interaction with sites of memory storage such as the hippocampus, striatum and cerebral cortex. However, how memory is strengthened at sites of storage is unknown. Tones that acquire behavioral importance gain representation in the primary auditory cortex (A1) that satisfies the cardinal criteria for a cortical memory trace. We hypothesized that the amygdala can enhance memory consolidation by facilitating such cortical representations. We have shown that pairing a tone with stimulation of the basolateral amygdala in anesthetized preparations shifts the tuning of cells in A1 to respond best to the paired frequency. We report here that amygdala stimulation induces representational plasticity within the waking animal. Frequency response areas (multi-unit responses to randomly presented combinations of pure tone frequencies and intensities) were obtained from adult male rats prior to and up to 22 days after a single 60 trials discrimination training session: CS+ tone paired with amygdala stimulation, CS- tone unpaired. We found two populations of cells in A1 within each animal that exhibited specific plasticity. One developed a CS+ tuning shift of 24 h duration. The second developed enhanced responding to the CS+ beginning 24 h after training, continued to grow for 7 days and maintained for at least 22 days. These findings show that the amygdala can enhance cortical memory traces via dual mechanisms. Moreover, their different time courses within an animal are compatible with the involvement of transient tuning shifts in long-term representational plasticity in the cerebral cortex. They may also provide a mechanism for increased behavioral memory strength that is a function of increased signal representation in A1.

Support: NIMH MH 12526
Automatic phoneme categorization in the dorsal auditory pathway

AUTHOR BLOCK: *M. A. CHEVILLET, X. JIANG, J. P. RAUSCHECKER, M. RIESENHUBER;
Neurosci., Georgetown Univ., WASHINGTON, DC

Abstract:
Current theories of speech perception in the human brain include two cortical processing streams, a ventral and a dorsal stream. The ventral stream is generally believed to mediate acoustic decoding of the speech signal and, ultimately, linking sounds to meanings. The function of the dorsal stream is less well understood, however, and has been a matter of some debate. It has recently been postulated that the dorsal stream may play a key role in sensorimotor integration linking speech sounds to motor articulations (Rauschecker & Scott, 2009). To further investigate this hypothesis we probed the neural representations of phonemes in the human brain using a novel fMRI rapid adaptation (fMRI-RA) paradigm. In fMRI-RA, two stimuli are presented in each trial, and the resulting BOLD-signal is thought to reflect the dissimilarity between neuronal activation patterns for the two stimuli. Subjects (N=14) were asked to perform a demanding distracter task (offset detection, mean accuracy = 71%) on pairs of speech sounds, created by auditory morphing between digitized natural /da/ and /ga/ utterances. By pairing speech sounds of comparable acoustic dissimilarity from either the same or a different phonetic category we could dissociate neuronal selectivity for acoustic features vs. phonetic categories. In left anterior middle temporal gyrus (aMTG) and left posterior superior temporal gyrus (pSTG) we found adaptation for same-stimulus pairs vs. acoustically different pairs from the same category (p<0.005), and no difference between acoustically different pairs from either the same or different categories (p>0.19). In left premotor cortex (PMC, BA6/9), we found adaptation for sound pairs from the same category compared to pairs from different categories, irrespective of acoustic similarity (p<0.005). A functional correlation was also found between the PMC and pSTG ROIs (p<0.002) but not between the aMTG and PMC ROIs (p>0.87), controlling for correlations between aMTG and pSTG. These results support a model of speech processing in which a ventral stream maintains an acoustic representation of phonemes in aMTG that can be utilized in a task-dependent fashion akin to analogous findings in the visual domain (Jiang et. al., 2007), and an automatic "audition-for-action" processing hierarchy in the dorsal stream consisting of sensorimotor representations of speech sounds in pSTG and articulatory phoneme categories in PMC. Interestingly, category-selectivity in PMC correlated with behavioral ability to categorize phonemes outside the scanner (p<0.02, r = 0.63), suggesting that dorsal stream phoneme representations may be recruited for speech categorization tasks.
Stimulus-specific adaptation measured in the guinea pig using magnetoencephalography G. B. Christianson(1), M. Chait(1), A. de Cheveigné(2), J. Linden(1,3)
(1) UCL Ear Institute, University College London, London, UK
(2) CNRS, Paris, France
(3) Department of Neuroscience, Physiology, and Pharmacology, University College London, London, UK

Using a newly developed magnetoencephalograph (MEG) for small animals, we have measured cortical responses to onsets and sound transitions in guinea pig and gerbil using tone complexes. The small-animal MEG system has 9 magnetometers placed in an 8x8 mm square array. An additional set of 3 magnetometers and one accelerometer are used to measure and suppress environmental noise. Sound is delivered using Etymotics transducers in either closed- or free-field conditions. Using this system, we have characterised some basic properties of sound-evoked MEG responses in rodents. Auditory onset responses occur with a latency of approximately 50 ms and last 300-400 ms, roughly half that observed in humans, while offset responses are extremely weak. We have also observed MEG responses consistent with stimulus specific adaptation (SSA). When short pips presented at a regular repetition rate were irregularly switched between two frequencies, a greater response was obtained for the first tone following a transition than for later tones. Consistent with previous physiology results, the underlying adaptation was extremely rapid and largely complete by the second pip in a sequence. In the long term, joint MEG and electrophysiology in the same animals will allow us to elucidate the neural basis of the MEG response, bridging the gap between human brain imaging and invasive animal electrophysiology.

This work was supported by the Biotechnology and Biological Sciences Research Council (BB/H006958/1), the Wellcome Trust (093292/Z/10/Z), Deafness Research UK, and the Programme CNRS - Neuroinformatique "Sounds to brain" (2009-2011).
Neural correlates of acoustic variability in conspecific vocalizations

K. L. CHRISTISON-LAGAY, S. BENNUR, Y. E. COHEN
Dept. of Otorhinolaryngology, Univ. of Pennsylvania, Sch. of Med., Philadelphia, PA

Abstract:
During the categorization of auditory objects, listeners ignore some sources of acoustic variability while paying attention to others, in order to extract the invariant characteristics of a given category. Vocalizations are a key mode of communication for both humans and non-human primates and thus may be among the most biologically salient of auditory objects. However, just as with other categories of auditory objects, there is variance between exemplars within a class of vocalizations. We hypothesized that the amount and type of acoustic variability within and between classes of vocalizations may be a fundamental property by which listeners categorize vocalizations and discriminate between vocalizations. To test this hypothesis, neural activity in the lateral belt of the auditory cortex was recorded; this region of the cortex is thought to be involved in processing auditory object identity. Neural activity was recorded while monkeys participated in a vocalization-discrimination task in which exemplars from two classes of vocalizations (coos and grunts) were presented. The monkeys released a lever to report when a “test” exemplar was different than the “reference” exemplar. We then correlated the acoustic variance between vocalization exemplars with the monkeys’ behavioral performance and with the neural activity. We found that the monkeys’ behavioral reports, neural activity and acoustic variability between the vocalization exemplars weakly covaried. These data suggest that this region is an intermediate region, representing both the acoustics of a stimulus and object identity. These data are consistent with a role for this region of the auditory cortex in vocalization perception. Furthermore, these findings reinforce the hypothesis that the pathway between the auditory cortex and the ventrolateral prefrontal cortex is an important functional pathway for processing auditory-object identity.
The basal ganglia in perceptual timing: timing performance in Multiple System Atrophy and Huntington’s Disease

AUTHOR BLOCK *T. E. COPE¹, M. GRUBE¹, B. SINGH³, D. J. BURN², T. D. GRIFFITHS¹;
¹Medical School, ²Institute for Ageing and Health, Newcastle University, Newcastle-upon-Tyne, United Kingdom; ³Walkergate Park Centre For Neurorehabilitation and Neuropsychiatry, Newcastle-upon-Tyne, United Kingdom;

Abstract:
This work explores the role of the basal ganglia in the timing of auditory events. Previous studies of the perception of acoustically presented time intervals and rhythms at and above the millisecond level implicate areas distinct from auditory cortex (Cope et al., 2011). Recent studies investigating the effects of cerebellar transcranial magnetic stimulation (Grube et al., 2010) and cerebellar degeneration (Grube et al., 2010) imply an obligatory role for the cerebellum in absolute, duration-based timing of perceptual events, but not in relative timing based on a regular beat. Functional imaging studies further support the cerebellum as a major constituent of a central timing network, but demonstrate preferential activation of the striatum for relative, beat-based timing (Grahn & Brett, 2007; Teki et al., 2011).

In the present work, six perceptual timing tasks were carried out to assess absolute duration discrimination for sub and supra-second time intervals, detection of beat-based regularity and irregularity, detection of a delay within an isochronous sequence, and the discrimination of sequences with metrical structure. Test groups comprised of individuals with Multiple System Atrophy, a disorder in which similar pathology can produce clinical deficits associated with dysfunction of the cerebellum (MSA-C) or striatum (MSA-P (Burn & Jaros, 2001)), and early symptomatic Huntington’s disease (HD), and a control group of individuals with chronic autoimmune peripheral neuropathy.

All patient groups exhibited impairments in performance in comparison with the control group for all of the timing tasks, and severity of impairment was overall associated with disease of longer duration. The data support an obligatory role for the striatum in all timing tasks, both absolute and relative, and call into question models of a brain timing network based upon independent modules.

References:
Behavioral state modulates accuracy of stimulus reconstruction from single-trial neural activity in auditory cortex

Stephen V. David, Bernhard Englitz, Jonathan B. Fritz, Shihab A. Shamma
Institute for Systems Research, University of Maryland, College Park

Sensory behavior requires real-time readout of task-relevant information from the neural population response to a single stimulus presentation. We used linear decoding methods to investigate: (1) how the accuracy of sensory representations is affected by correlated neural noise and (2) how representations change as specific stimulus features become behaviorally relevant.

To study the effects of noise correlations, we recorded the simultaneous activity of groups of neurons (n=66 recordings, 2-10 neurons per experiment) in ferret primary auditory cortex (A1) during passive listening to a bandpass noise stimulus modulated by a naturalistic temporal envelope. We measured the accuracy with which the stimulus envelope could be reconstructed from single-trial responses using a linear decoder and compared this result to one obtained after shuffling responses of individual neurons across trials, thereby removing noise correlations. We found that shuffling systematically improved reconstruction accuracy (mean correlation between actual and reconstructed envelope: r=0.47 non-shuffled, 0.51 shuffled), reflecting an average decrease of about 15% in explained stimulus variance due to noise correlations. A second-order decoder that explicitly accounted for pair-wise correlated activity showed no improvement over the linear decoder, indicating little or no synergistic stimulus information in correlated A1 activity beyond what can be measured with a linear decoder.

To study the effects of behavior on the population code, we measured the activity of neural populations in A1 during a task that required the discrimination of amplitude modulated tones (12 Hz) from unmodulated pure tones, both with random carrier frequency (100 Hz to 8KHz). A linear classifier was used to decode the modulation rate and carrier frequency during behavior and during passive listening to the same sounds. Preliminary data from one animal (n=22 recordings) revealed that during the behavior, modulation rate could be classified with about 8% greater accuracy while carrier frequency could be classified with slightly worse accuracy. These results suggest that task-driven attentional modulation in A1 serves both to enhance the discriminability of task-relevant stimulus features and to diminish the discriminability of task-irrelevant features. Taken together, these findings demonstrate that intrinsic brain activity can influence the accuracy of representations in A1 in ways that emphasize task relevant features.
Modeling of harmonic sensitive neurons in the primary auditory cortex of marmoset monkeys

AUTHOR BLOCK: *E. DEKEL, L. FENG, X. WANG, K. ZHANG;
Johns Hopkins Univ., Baltimore, MD

Abstract:
Neurons that are sensitive to harmonic complex tones have been shown to exist in central auditory structures from the inferior colliculus to the auditory cortex. Our goal in this project was to investigate whether a computational model comprised of excitatory and inhibitory units can account for the high-level complex neural responses to various simple and complex stimuli in the primary auditory cortex, using data recorded from harmonic-sensitive neurons in the auditory cortex of awake marmosets (Callithrix Jacchus). The stimuli presented to these neurons included pure tones, two-tone combinations, random harmonic tone stacks and jitter. Three models with increasing level of complexity were designed to account for the observed cortical responses, representing the neural connections in different stages of the auditory processing hierarchy. All of the models included the combination of inhibitory and excitatory neurons, modified from previously suggested models of A1. Models 1 and 2 included a simple feed-forward network with three excitatory neurons and two inhibitory neurons, all converging to a final output, with (Model 2) and without (Model 1) lateral connections between them. These two models consider the thalamus as a relay and represent the cortical layers alone. Model 3 simulates the auditory processing from the IC, through the thalamus and into A1. The recorded data from each neuron was divided to optimization and validation data sets, and the parameters of each model were fitted based on the optimization data set using exhaustive search optimization of the mean square error (MSE) between the measured and predicted responses. The model's predictive capability was measured using the unseen validation data set, and the models were compared using the MSE and the Bayesian information criterion (BIC), which also penalizes for increased model complexity. The fitting results suggest two main points: 1) Complex models (such as Model 3) fit the data better than the simple ones (Models 1 and 2), but their large number of parameters gives them a higher BIC score. Also, for the simple stimuli presented here, these complex models do not generalize as well as the simple ones, probably due to over-fitting of noise. 2) A reoccurring solution to the fitted models suggest that lateral connections between the inhibitory and excitatory units support a co-tuned relationship between their weights (similar mean and variance), forming a subunit that may possibly underlie the observed neural responses to harmonic sound structures.
Effects of congruent and incongruent face and vocalization pairs on neurons in ventral prefrontal cortex.
M. M. Diehl, M. D. Diltz, L. M. Romanski.

Social communication relies on the integration of both auditory and visual information that are present in faces and vocalizations. Evidence has suggested that the integration of multisensory information enhances the perception of social communication stimuli compared to the processing of stimuli from a single modality in both humans and non-human primates. The ventral frontal lobes are one of several brain regions involved in processing social communication information. Our previous studies have demonstrated that neurons in ventrolateral prefrontal cortex (VLPFC) respond to conspecific vocalizations and their accompanying facial gestures presented separately and combined. In the present study, we asked whether neurons in VLPFC detect an incongruent cross-modal communication stimulus. We recorded single-cell activity during the presentation of incongruent audiovisual vocalization movies. Stimuli included video clips of non-human primates engaging in affiliative and agonistic vocalizations. Auditory and visual components of the stimuli were presented simultaneously and separately to assess unisensory and multisensory responses in VLPFC. The simultaneous presentation involved both congruent and incongruent (mismatched) presentations. Examples of the incongruent stimuli included an agonistic facial gesture paired with an affiliative vocalization and vice-versa. Recordings in VLPFC showed that while many cells are multisensory and respond to faces, vocalizations, or their combination, a subset of neurons show a significant change in neuronal activity during incongruent cross-modal vocalizations. These results show VLPFC neurons discriminate congruent and incongruent face-vocalization stimuli. Moreover, it confirms the role of the ventrolateral prefrontal cortex in the processing and integration of multisensory communication information.
Spectro-temporal Neural Coding of Speech in Human Auditory Cortex

Nai Ding, Jonathan Z. Simon

University of Maryland, College Park

We study the neural representation of continuous speech in human auditory cortex. The Magnetoecephalographic (MEG) response was recorded from human subjects listening to a spoken narrative, either in quiet or in the presence of interfering speech. For the speech signal presented alone, it is demonstrated that the ongoing neural activity in auditory cortex precisely follows the slow temporal modulations (< 10 Hz) of speech in a broad spectral region between 400 Hz and 2 kHz. This temporal neural code observed by MEG is sufficiently faithful to reconstruct the envelope of speech.

To examine the robustness and attentional modulation of this neural code, we presented two spoken narratives simultaneously and instructed the listeners to focus on only of them. The two spoken narratives are either from the same speaker but presented to different ears, or from different speakers and presented to the same ear. In auditory cortex, we find the neural representation of the speech attended to is substantially stronger than that of the unattended speech. This attentional modulation is significant even during the subjects’ first exposure to the spoken narratives. To summarize, these results demonstrate that human auditory cortex precisely represents the slow temporal modulations of speech using a temporal code. Moreover, concurrent speech signals are segregated in auditory cortex and are represented differently under the modulation of top-down attention.
Intracranial Neural Correlates of Auditory Perceptual Awareness

Andrew R. Dykstra¹,², Eric Halgren³, Alexander Gutschalk⁴, Alexandra J. Golby⁵, Emad Eskandar⁶, Sydney S. Cash²

¹Program in Speech and Hearing Bioscience and Technology, Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA
²Cortical Physiology Laboratory, Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA
³Departments of Radiology and Neurosciences, University of California San Diego, San Diego, CA
⁴Department of Radiology, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany
⁵Department of Neurosurgery, Brigham and Women's Hospital and Harvard Medical School, Boston, MA
⁶Department of Neurosurgery, Massachusetts General Hospital and Harvard Medical School, Boston, MA

In complex acoustic environments, our ability to perceive specific sounds of interest is thought to be limited by information-processing bottlenecks in the central auditory system. The neural activity underlying the conscious perception of target sounds in such listening situations has not been well characterized. We recorded the intracranial EEG from three neurosurgical patients with epilepsy who reported when they began to detect regularly-repeating target tones amongst random multi-tone maskers. Neural responses including evoked potentials and high-gamma activity were compared between conditions in which targets were detected or undetected. Compared to undetected targets, detected targets elicited early (~50-200 ms) activity in the posterior superior temporal gyrus (i.e. auditory cortex) as well as a broad, long-latency (~300-600 ms) response that was widespread, including auditory cortex as well as ventrolateral prefrontal and anterolateral temporal cortices. The results demonstrate that the neural activity associated with the perceptual awareness of specific target sounds in complex settings engages diverse brain areas, including early sensory cortex as well as supramodal areas known to be involved in attentional selection and target detection.
Spike-timing reliability and cross correlations are enhanced by a partial blockage of GABAA inhibitions in the guinea pig auditory cortex.

Abstract:
Over the last decade, several laboratories have provided evidence suggesting that a spike timing code operates in auditory cortex (ACx) for discriminating natural stimuli. One intriguing question is whether the spike timing reliability observed at the cortical level depends upon the interplay between excitations and inhibitions converging on a given cortical cell, or if it is generated subcortically and passively transmitted to the ACx. To answer this question, we recorded neuronal activity in ACx of urethane anesthetized guinea pigs before and after pharmacological blockage of intracortical inhibitions. Multiunit activity was collected by 16 electrodes arrays positioned in the tonotopic field A1. Tuning curves were quantified by presenting gamma tones from 0.1-36kHz at 75-45dB, then responses to conspecific and hererospecific vocalizations were tested. The protocol was repeated immediately after the application of GABAA or GABAB antagonists (Gabazine, Saclofen and CGP55485 all at 10µM, 4-minutes pulse application). Immediately after application, Gabazine (GBZ) significantly increased the tuning width, the response strength (to tones and to vocalizations) and surprisingly it also increased the spike timing reliability and cross-correlation levels (in evoked and spontaneous activities). The effects on tuning width and response strength maintained 1h post application whereas the spike timing reliability and the cross-correlation levels returned to control values. Application of saline, Saclofen or CGP did not influence the responses to artificial and natural sounds, the spike timing reliability and the strength of the cross-correlations.
These results indicate that GABAA inhibitions partly control the responses strength to artificial and natural stimuli as well as the timing of neuronal discharges and of neuronal interactions.

Supported by ANR “Hearing Loss”, the FRC and the CNRS NeuroInformatique program.

1937 caractères.

Keywords: Gabazine, Saclofen, Spike timing, Cross-correlations, Synchronization
A study on rat vocal interactions and syntax

M. ELIAS COSTA 1, A. M. WIDMER 2, M. SIGMAN 1, M. N. GEFFEN 3, D. A. LAPLAGNE 2;

1 Dept. de Física, FCEyN, Univ. de Buenos Aires, Buenos Aires, Argentina;
2 Shelby White and Leon Levy Ctr. for Mind Brain and Behavior, Rockefeller Univ.,
New York, NY;
3 Dept. of Otorhinolaryngology, Univ. of Pennsylvania Sch. of Med., Philadelphia,
PA

Abstract:
Rats and mice produce vocalizations that span the frequency range from the sonic
to the ultrasonic. Sonic calls are scarce and produced only in extreme scenarios
like predator encounters. In the rat, emission of prolonged 22 kHz calls
correlates with aversive and fearful conditions and has been shown to convey
alert to listeners. Above 40 kHz, rats produce a rich repertoire of vocalizations
spanning a wide range of frequency, frequency modulation and duration. Semantic
studies on these ultrasonic vocalizations (USVs) have shown a correlation with
positive situations like satiety, reward and social interaction, although the
fine semantics of call types within this family remains elusive. On the syntax
side, few studies have shown that mice can sequence USVs in neither random nor
completely stereotyped motifs, while syntax remains unexplored in rats.
Furthermore, despite the social role of vocal communication, no study has
analyzed the dynamics of vocal interactions between rodents, mainly due to the
difficulty of assigning calls to individuals in a common acoustical environment.
We have developed a setup for recording USVs from pairs of rats. The arena is
split in two, so that rats remain separate from each other but have perfect
acoustical connection. Individual calls are recorded from two directional
microphones hanging above and can be assigned unambiguously to their emitting rat
by their difference in power on each mic. We have developed a full set of
algorithms to automatically detect, assign and classify the calls. Video tracking
is also implemented to correlate vocal to other behaviors.
This simple behavioral arena promotes high rates of USV production, averaging up
to 2 calls per second per rat and amounting to over 100,000 so far recorded. Call
production has a characteristic timescale, most of them coming in bouts of 7 Hz.
Rats vocalize more when paired with another than when alone, and strongly tend to
do so towards the center of the arena, facing their companion. We present
analysis on both the sequencing of calls within single rat speeches and dynamics
that emerge from their vocal interactions.
The rat is a potentially powerful animal model to study the neurobiology of
speech production and perception given their highly social ethology and
amenability to physiological and behavioral studies in the lab.
Title:
Population Representation of Shepard Tones in the Auditory Cortex of the Awake Ferret

Authors:

Affiliation:
Neural Systems Lab, University of Maryland, College Park, USA

Abstract:
Perceptual effects on neural activity can best be studied if the same physical stimulus gives rise to different percepts, e.g. for bistable or illusory stimuli. Shepard tones are complex acoustical stimuli consisting of an octave-spaced stack of pure tones. A sequence of two Shepard tones separated in frequency by 6 semitones is bistably perceived as either ascending or descending in frequency, the so-called 'tritone paradox'. A sequence of Shepard tones played with the base pitch of tone moving upwards or downwards creates the auditory illusion of a tone that continually ascends or descends in pitch. Shepard tones are thus useful acoustic stimuli to study auditory perception.

To explore their neural representation, we presented various sequences of non-ambiguous and ambiguous Shepard pairs to passively listening and behaving ferrets and recorded single unit neuronal responses from >350 neurons from 5 ferrets in primary auditory cortex A1 and secondary auditory regions PPF and PSF. Neurons responded to Shepard tones with a wide variety of responses based on their basic response properties. By applying a range of dimensionality reduction techniques, we obtained a low-dimensional representation of the Shepard tones from the neuronal responses. Shepard tones with different base pitch mapped to a circular structure thus capturing the periodic nature of the Shepard tone stimulus. We presented the ambiguous Shepard pair in different auditory contexts which have been demonstrated to influence the perception of the pair (Chambers et al. in preparation) and compared the neural responses for both upward and downward contexts. Results from passive and active conditions will be presented.
Neural coding of harmonic complex tones in auditory cortex of awake marmosets

Lei Feng & Xiaoqin Wang

Laboratory of Auditory Neurophysiology, Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland

Many natural and man-made sounds, such as animal vocalizations, human speech and music are harmonically structured. Harmonicity is an important cue used by the auditory system to group together the harmonically-related elements of one sound source to form a single percept and also to segregate sounds from different sources, which likely have harmonically-unrelated spectral components. One simple example is that a mistuned harmonic is easily heard as a ‘pop-out’ from the remaining harmonics. However, the neural mechanisms for grouping and segregation based on harmonicity are largely unknown. In the present study, we systematically examined the representations of harmonic and mistuned complex tones by single neurons in auditory cortex of marmosets (Callithrix jacchus)). We found a subpopulation of neurons in primary auditory cortex whose responses were shaped by harmonicity in spectral context. Neural responses to a tone at characteristic frequency (CF) could be significantly enhanced when it was presented simultaneously with other harmonically related tones even though these context tones did not induce any responses in isolation. When the CF tone was mistuned from other harmonics, the firing rate decreased. Moreover, the CFs of these neurons were closely associated with the frequency components of marmoset vocalizations. Our observations suggest that a subset of neurons in auditory cortex function to extract particular harmonic spectral structures embedded in complex sounds. Such functional properties could result from the interaction between excitatory and inhibitory receptive field components of a neuron or a network of neurons in auditory cortex or its inputs.
Comparison of task related plasticity in ferret primary and secondary auditory cortex.  
Jonathan B. Fritz¹, Stephen David⁴, Serin Atiani², Michael Locastro¹, Diego Elgueda⁴,  
Susanne Radtke-Schuller³, Shihab A. Shamma¹

¹Institute for Systems Research, University of Maryland, College Park, MD, USA; ²BRAMS,  
University of Montreal, Montreal, Canada; ³Division of Neurobiology, Biocenter of Ludwig  
Maximilians University, Munich, Germany.

Rapid task-related receptive field plasticity alters the tuning properties of neurons in  
primary auditory cortex (A1) in a task-specific fashion that enhances their ability to  
encode salient stimuli in the current task (Fritz et al., 2003). We investigated task-related  
receptive field and response plasticity in two non-primary, tonotopically organized  
cortical areas (PPF and PSF) in the posterior ectosylvian gyrus (PEG) of the ferret  
(Bizley et al., 2005, Nelken et al., 2004) and compared the properties of PEG and A1  
plasticity in identical task conditions. Previous studies have reported greater receptive  
field plasticity in secondary auditory areas (AII in the cat vs A1) induced by classical  
conditioning (Diamond and Weinberger, 1984) and increased frequency selectivity in  
secondary auditory areas (VPAF vs A1 in the rat) induced by paired tone-nucleus basalis  
stimulation (Puckett et al., 2007). We trained six ferrets on a conditioned avoidance  
auditory task that required them to distinguish a set of rippled noise or band pass noise  
stimuli (reference stimuli) from pure tones (target stimuli). We physiologically mapped  
the tonotopic gradient in the middle and posterior ectosylvian gyus and found mirror  
image tonotopic maps, corresponding to A1 and the two tonotopic areas of PEG. We  
recorded separately or simultaneously from single units in A1 (120 neurons) and PEG  
(250 neurons) during two behavioral states: while the animals were in a quiescent state of  
passive listening, and while in an active state during task performance. We pooled  
responses of all neurons in PPF and PSF. Receptive fields in these areas exhibited  
changes reflecting the target and reference stimuli in a similar manner to A1. However,  
unlike A1, we observed robust task-related changes in firing rates and response dynamics  
in PEG. Specifically, responses to behavioral target vs reference stimuli diverged  
substantially during the active state for many neurons, presumably allowing for greater  
discrimination in these cells between the two classes of acoustic stimuli during active
listening vs quiescent (non-task) listening. Such response modulation, observed at a single cell and population level, may be pivotal for categorical encoding of attended auditory objects, previously described in secondary auditory cortex (Tsunada et al., 2011) and prefrontal cortex (Cohen et al., 2006; Fritz et al., 2010). Insight into the differential types and magnitude of task-related plasticity in secondary auditory cortical areas may clarify the influence of auditory attention on A1 and these multiple non-primary auditory areas and their functional contributions to acoustic representation and perception.
Temporal dynamics of the tonotopic maps in awake primates

M. FUKUSHIMA, R.C. SAUNDERS, D.A. LEOPOLD, M. MISHKIN, B.B. AVERBECK

Lab. Neuropsychology, NIMH/NIH, Bethesda, MD.

Abstract:

The macaque auditory cortex is tonotopically organized to represent sound frequency in multiple regions on the supratemporal plane (STP). While neural responses in these areas have been studied in detail, the spatiotemporal dynamics of these maps in response to acoustic stimuli are not well understood in awake monkeys. To examine this issue, we designed a micro-electrocorticographic (µECoG) array to record field potentials simultaneously from multiple areas of macaque auditory cortex. Each µECoG array has 32 recording sites on a 3x7mm grid with 1mm spacing. We chronically implanted four such arrays in each of two monkeys, allowing for 128 parallel measurements from the surface of the STP as well as the caudal superior temporal gyrus. First we examined the auditory frequency tuning of auditory evoked potentials with pure tone stimuli at multiple conventional frequency ranges of the field potentials. This analysis revealed tonotopic maps that reversed frequency direction at putative areal boundaries. The smoothest tonotopic maps were obtained from the high-gamma band. Next, we estimated when each site showed significant discrimination among different stimulus frequencies by evaluating the high-gamma power in a moving window with an ANOVA. We found that the onset time of the discrimination increased along the caudal-to-rostral as well as the medial-to-lateral axes, consistent with the hypothesis that auditory information is serially processed in those two directions in parallel. Taking further advantage of the simultaneous recordings we next evaluated the undriven resting activity by using principal components analysis. We found that the first several principal components, which represented the dominant spatial modes in the spontaneous activity, significantly correlated with the tonotopic map. This suggests that the spontaneous fluctuation in the activity of the auditory cortex reflects the functional architecture of the network.
Neuronal adaptation in the awake rat auditory cortex depends on the spectrotemporal features of structured sound stimuli

B.H. GAESE, K. SCHMALE, C. KLEIN
Inst. Cell Biol. and Neurosci., Goethe University Frankfurt, Frankfurt/Main, Germany

Neurons in the auditory cortex code for the physical properties of a stimulus by increasing or decreasing their firing rate. Besides physical parameters describing an acoustic object the context in which it appears also plays an important role. Detecting changes in the surroundings that might indicate novel events is of strong behavioral relevance. In the auditory domain the phenomenon of mismatch negativity (MMN) has been related to such structuring processes. However, the EEG/MEG experiments in humans that mainly describe this phenomenon left the neural mechanism generating MMN unclear. Stimulus-specific adaptation (SSA) of neurons in the auditory pathway has been suggested to contribute to MMN at the cellular level. Like in MMN experiments an infrequent stimulus (deviant) is presented in a series of repetitive stimuli (standard) and their responses are compared. While most studies focused on pure tones as stimuli we also used upward and downward frequency modulated tones (FM), structured stimuli containing tonal and noise elements, and periodically modulated stimuli to investigate SSA with more complex tones. Several of those stimulus types induced SSA but the underlying mechanisms seemed to differ. While for simple stimuli adaptation was found in the short-latency response components, more long-latency response components showed adaptive behavior for structured acoustic signals. An additional pattern was, that the more temporally structured the stimuli were, the less stimulus-specific adaptation could be induced. This indicates a more sophisticated pattern of underlying mechanisms and a presumably not so straight-forward general functional implication.
Modulation of slow oscillations to the sound coding in auditory thalamus
Lixia Gao, Jufang He

CAS-Hong Kong Joint Research Laboratory for Visio-Auditory Integration
Department of Rehabilitation Sciences, Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong

The natural fluctuations of membrane potential between a hyperpolarized DOWN state and a depolarized UP state are observed from neurons of cortex and thalamus during slow wave sleep and in the anesthetized or even waking animals. However, the functional impact of these membrane potential fluctuations is still unknown. In the present study, we investigated whether and how the slow oscillations of membrane potentials modulate the sound coding in medial geniculate body (MGB), with in vivo intracellular recordings of anesthetized guinea pig. We artificially manipulated the resting membrane potential with current injection to simulate the slow oscillation during slow-wave sleep. We found that (1) acoustic responses at UP and DOWN states were different in the MGB. (2) Responses at UP and DOWN state depended on the sound frequency and the sound intensity and (3) the threshold of the neuronal responses to pure tone was different at UP and DOWN states. (4) The neuronal responses in the MGB increased after the simulated Up and DOWN membrane potential fluctuations. The results suggest that slow-wave sleep positively facilitate the sound encoding in the awake animal.

(Supported by Hong Kong Research Grants Council: PolyU5614/10M)
Synaptic inputs to large tectothalamic projection neurons of the mouse inferior colliculus

AUTHOR BLOCK: H.-R. A. P. GEIS, *G. G. BORST, Prof; Erasmus MC, Rotterdam, Netherlands

Abstract:
Neurons in the inferior colliculus (IC) show a remarkable diversity in their responses to sound, but it has proven difficult to relate the different sound responses to their morphology. Large cells, which are found in all subdivisions of the IC, may form a distinct subclass, as they have been shown to be GABAergic, to have vesicular glutamate transporter 2 containing terminals on their somata and to project to the thalamus. To test whether these cells indeed have distinct electrophysiological properties or sound responses, we recorded intracellular responses to tone stimulation in large neurons of the dorsal cortex of the mouse inferior colliculus, which we targeted under two photon guidance. Large cells typically had a relatively low input resistance. The majority of these cells received short latency excitatory inputs and fired short latency action potentials, in agreement with the presence of prominent somatic glutamatergic inputs. The excitatory inputs were often followed by long latency inhibitory postsynaptic potentials. In four cells it was possible to reconstruct their ascending axon following labeling with biocytin. Each of the cells projected to the thalamus, but also branched off large collaterals while passing through the brachium of the inferior colliculus. Our data support the notion that owing to their somatic glutamatergic inputs, GABAergic tectothalamic projection neurons can generate short latency feed-forward inhibition, which affects not only the thalamus, but also other nuclei downstream from the IC. The presence of delayed inhibitory postsynaptic potentials in these neurons might help to keep this inhibition both brief and well-timed.
Structural correlates of auditory and phonological skill in school children

Manon Grube, Sukhbinder Kumar, Freya Cooper, Timothy D Griffiths.

Newcastle University, Auditory Group, Medical School, Institute of Neuroscience, Framlington Place, Newcastle-upon-Tyne, NE2 4HH.

This work seeks a structural correlate of auditory skill & language ability in school children (age 13) who have previously undergone detailed behavioral testing (at age 11). Behavioural testing was run in one year (n=238) of the St. Thomas More School (Newcastle, UK) and included 12 auditory tasks (4 pitch, 4 rhythm, 4 modulation) at a range of levels of complexity, 6 standard tasks of phonological skill, and standard intelligence tests. The behavioral data demonstrate a significant relationship between auditory and phonological skill in the larger group with a prominent relationship between pitch sequences or rhythm and phonological skill. Structural MRI has been carried out at 3T on a subset of children (n=42) from across the range of task performance and intellectual ability. The subset shows the same relationship between auditory and language skill (Spearman’s rho = 0.52, p <.01) as found in the broader population. Voxel-based morphometry implemented in SPM 8 was used to seek correlation between grey matter density and auditory & phonological skill, taking into account non-verbal intelligence. Significant correlation is demonstrated in the region of the left intraparietal sulcus and angular gyrus between grey matter density and a composite measure of auditory skill (Pearson r = 0.47, p <.05). A relationship between grey matter density and phonological skill (Pearson r = 0.55, p <.01) was demonstrated in the same areas. Further analysis is being carried out to assess in more detail the relationship between specific auditory and phonological skill and brain structure. The preliminary analysis so far would support a structural basis for either domain and in addition possibly also for the link between the two: auditory and language skills.
Title: Noise-invariant representation of speaker identity in human auditory cortex.

Authors
Lars Hausfeld, Federico De Martino, Milene Bonte, Elia Formisano

Abstract
Listeners effortlessly identify familiar voices independently of speech content and in the presence of background noise. In this study we examined the neural representations at the basis of speaker identification in the context of natural auditory scene using fMRI and a decoding approach (similar to [1]).

Stimuli included vocal sounds of one of three speakers within one of three background conditions. In particular, 60 short non-linguistic vocalizations (≤ 1s) were recorded from 3 Dutch speakers. Noise conditions were created using no background sound, white noise or environmental noise recordings (e.g. cafeteria, rain). After familiarization with the voices, fMRI was performed with stimuli presented to participants in silent gaps between acquisitions. During scanning subjects performed a speaker identification task.

Behavioral results indicate that participants successfully identified the speakers’ voices for all background conditions. We decoded speakers’ identity based on cortical activity in superior temporal cortex using support vector machines. The classifications were significantly above chance both when training and testing was performed within a single background condition and when all background conditions were considered together. Temporal regions relevant for accurate speaker classification were found bilaterally in anterior-lateral Heschl’s gyrus, anterior STG and posterior STS/MTG and overlapped for different types of background.

Our results suggest that neuronal populations in early (HG) and higher order auditory regions (STG, STS, MTG) encode representations of speaker identity, which are invariant to content and noise. These representations may be pivotal to the identification of a familiar voice in a noisy environment.

Crossmodal thalamocortical connections of the primary auditory, somatosensory and visual cortex in the Mongolian gerbil

Julia Henschke, Henning Scheich, Eike Budinger

Leibniz Institute for Neurobiology Magdeburg, Germany, Dept. Auditory Learning and Speech

It has long been the prevailing view that primary sensory cortices like the primary auditory (A1), somatosensory (S1) and visual field (V1) are unimodal, i.e. are dominated by neurons responding only to their own sensory modality. However, very recent studies have shown there are also neurons in A1, S1 and V1, which respond to stimuli of other sensory modalities. This raises the question about the anatomical pathways how the multisensory information is forwarded to these cortical areas?

Here, we used the retrograde transport of two sensitive and fluorescent neuronal tracers, namely fluorescein-labeled dextranamine (FDA) and tetramethylrhodamine-labeled dextranamine (TMRDA), in order to examine the thalamocortical afferents of A1, S1 and V1 in the Mongolian gerbil (Meriones unguiculatus), a commonly used animal model in multisensory research (Budinger & Scheich, Hear Res 258:16-27, 2009). Following simultaneous injections of FDA and TMRDA into these areas (in combination: A1/S1, V1/S1, A1/V1) we found overlapping clusters of single FDA- and TMRDA-labeled as well as double-labeled cell bodies in several sensory thalamic nuclei. This indicates that there are individual thalamic nuclei as well as individual neurons, which project in a divergent manner to more than one primary sensory area. Overlapping clusters were found in several nuclei of the medial geniculate body (medial division, dorsal division, marginal zone and suprageniculate nucleus), ventral thalamus (ventral posterolateral, ventrolateral and ventral posteromedial nucleus), lateral thalamus (lateral posterior and laterodorsal nucleus) and posterior thalamic nuclear group. Double-labeled cells were found in the medial geniculate body (medial division, dorsal division and suprageniculate nucleus), posterior thalamic nuclear group and in the brachium of the inferior colliculus.

In summary, approx. 4% of the retrogradely labeled cells in the thalamus project not only to the primary sensory cortex of their own but also to cortices of other modalities. Although sparse, these thalamic afferents provide one possible anatomical basis for the integration of multisensory information at a very early cortical level.
Intracellular correlates of stimulus-specific adaptation

I. HERSHENHOREN, N. TAASEH, A. YARON, *I. NELKEN;
Hebrew Univ., Jerusalem, Israel

Abstract: Stimulus-specific adaptation (SSA) is the reduction in response to a common stimulus which does not, or only partially, generalizes to other, even rather similar, stimuli. SSA has been studied mostly in the auditory system. It is strong and widespread in primary auditory cortex (A1) of rats, but is rather weak or absent in the main input station to A1, the ventral division of the medial geniculate body. In order to study the transformation of SSA from the input to a neuron to its output, we recorded intracellularly in A1 neurons. Recordings have been performed using sharp electrodes, and lasted for up to 780 minutes. The responses to rare sounds were recorded in a number of different contexts, including oddball sequences, Deviant-alone sequences in which most sound presentations were replaced by silence, and sequences composed of many different sounds each of which was rare. SSA was found both in subthreshold membrane potential fluctuations and in the spiking responses of A1 neurons. By comparing the responses to the same rare sounds in multiple contexts, we can show that cortical SSA cannot be fully explained by adaptation in narrow frequency channels. SSA for changes in frequency was large at frequency differences of 44%, and clearly present with tones as close to each other as 4%, near the perceptual difference limen for frequency in rats. The amount of SSA was, however, only partially correlated between the membrane potential and the spiking output. In particular, some neurons had rather weak SSA at the level of their membrane potentials, but much stronger SSA when measuring their spiking output. Such differences may indicate the presence of processing mechanisms that amplify or create SSA de-novo in primary auditory cortex.
Frequency tuning in mouse auditory cortex depends on the sequencing of probe tones

K. Jannis Hildebrandt[1], Lucy A. Anderson[1], Maneesh Sahani[2] and Jennifer F. Linden[1,3]

[1] Ear Institute,
[2] Gatsby Computational Neuroscience Unit, and
[3] Department of Neuroscience, Physiology & Pharmacology,
University College London, London, UK

Frequency tuning of neurons in primary auditory cortex is not fixed, but can change with the behavioural task or the sensory context in which it is tested. For example, previous studies have suggested that estimates of frequency tuning might vary with the sequence and timing of tones used as probe stimuli, even when relatively large inter-tone intervals are used. Differences in frequency tuning emerging from interaction between successive probe tones may help us to understand how frequency tuning in the auditory pathway is shaped and altered by stimulus context. Here, we analysed frequency tuning of neurons in mouse auditory cortex using probe tones that were presented either in pseudo-random order (“randomly ordered tones”), or in sequential order by frequency at different sound levels (“sequentially ordered tones”). Inter-tone intervals were either 500 ms or 1 s; these intervals are as long or longer than the intervals typically used for characterising frequency-intensity response areas in most auditory cortex studies. We recorded both single-unit and multi-unit activity in the auditory cortex of ketamine-anaesthetised mice, and compared characteristic frequency (CF) as well as bandwidth and best frequency (BF) at different sound levels above threshold.

We observed no difference in CF between frequency-intensity response areas obtained using randomly versus sequentially ordered tones. However, tuning bandwidth was considerably broader when estimated from responses to randomly ordered tone sequences. Moreover, BFs at higher sound levels differed more from CF when estimated from responses to randomly ordered tone sequences than when estimated using sequentially presented tones. Thus, sequential ordering of tone frequencies produced not only narrower tuning curves but also more level-invariant coding of sound frequency. These results suggest that even when inter-tone intervals are long, successive presentation of similar tones refines frequency tuning in auditory cortex.

Acknowledgements: German Academic Exchange Service (DAAD), Wellcome Trust, Biotechnology and Biological Sciences Research Council, and Deafness Research UK. Thanks to G. Bjorn Christianson for technical advice.
Relationships among speech-evoked cortical oscillations, speech-evoked brainstem responses, and reading-related skills in children.

Hornickel J, Touny M, Escobedo Quiroz R, Kraus N

Oscillatory cortical activity, particularly in the gamma range, has been linked to a number of cognitive processes, such as memory, IQ, and temporal processing that may be important for developing reading skills. Cortical oscillations in the gamma range during infancy have been linked to later reading ability and language function\(^1\); however reading and language ability have not been concurrently assessed with gamma activity in response to speech sounds in older children. In addition, the relationships between cortical oscillatory activity and speech-evoked auditory brainstem responses have not previously been investigated. Speech-evoked auditory brainstem response measures, particularly measures of timing and harmonic representation important for distinguishing speech sounds, are correlated with reading ability and previous work has shown that brainstem response timing is linked to cortical response asymmetry in children\(^2,3\). Thus, we might predict that cortical oscillatory in the gamma range would be related to speech-evoked brainstem response measures in children. Here we explore these relationships among cortical activity, subcortical activity, and language-based learning in children.

Supported by NSF BCS-0921275, NSF SMA-1015614, and the Hugh Knowles Center, Northwestern University

1. Gou et al., Behav Brain Res, 2011
2. Hornickel et al., PNAS, 2009
3. Abrams et al., J Neurosci, 2006
Cross-modal Context-sensitive Responses to Combined Face-vocalization Stimuli in Ventrolateral Prefrontal Cortex of Non-human primates

*Jaewon Hwang¹ and Lizabeth M. Romanski² Rochester, NY; ¹Brain and Cognitive Sciences, University of Rochester, Rochester, NY; ²Neurobiology and Anatomy, University of Rochester, Rochester, NY

It has been shown that the prefrontal cortex is important for processing context information. For example, neurons in the dorsolateral prefrontal cortex are involved in cognitive context-dependent coding by responding differently to an identical stimulus according to the task situation. In the orbitofrontal cortex, there are neurons that code stimuli depending on the motivational context as well as neurons that represent motivational context information. Here we show evidence that neurons in the ventrolateral prefrontal cortex (VLPFC) also respond to a complex context defined by cross-modal stimuli. In our experiment, non-human primate subjects were trained to perform a non-match to sample task made with short movie clips which contained an auditory component (vocalization; An) and a visual component (face movie; Vn). The subjects were required to remember a movie clip presented during the sample period (A1V1 or A2V2) and detect a change that occurred in the video (A1V2 or A2V1; VIDEO non-match), audio (A2V1 or A1V2; AUDIO non-match) or a change that occurred in both the auditory and visual components (A2V2 or A1V1; AV non-match) during the subsequent non-match period. Across some conditions, the same stimuli were presented as non-match stimuli. However, depending on which stimulus was presented as the sample, identical stimuli could be an AUDIO non-match in some trials or a VIDEO non-match in other trials. For example, A1V2 was a VIDEO non-match when A1V1 was the sample and an AUDIO non-match when A2V2 was the sample. While the subjects performed the task, we recorded neural activity from the VLPFC. Although identical non-match stimuli were presented, some VLPFC neurons showed a greater response when the stimuli were presented as AUDIO non-match (or VIDEO non-match) and did not respond well when they were presented as VIDEO non-match (or AUDIO non-match). Interestingly most of these context-sensitive neurons responded to the AV non-match stimuli in a similar manner as to the VIDEO non-match stimuli. In other words, AUDIO non-match information in AV non-match stimuli did not make an effect on neural activity. Considering that the visual motion began earlier than the vocalization onset in our stimuli, our subjects did not need to evaluate the later-occurring auditory components of the movies once they detected a mismatch in the early visual component. Therefore, this ineffectiveness of AUDIO non-match information in AV non-match stimuli reflected the cognitive strategy of our subjects. Together, these results suggested that VLPFC may be involved in deciding which sensory information to use depending on cognitive requirements.
Our ability to adapt to different environments often depends on modifying the actions associated with a given stimulus. To study the neuronal mechanisms underlying flexible associations between sounds and actions, we developed a rapid reversal paradigm for rats (two-alternative choice task) in which the spectral content of an acoustic stimulus indicates the location of reward. Rats were trained to discriminate between two frequencies on each block of trials. In a first block, low-frequencies indicated reward on a left port, while middle-frequencies were associated with the right port. On a subsequent block, the middle-frequency was reversed to indicate left, and high-frequency stimuli were associated with reward on the right port.

Trained rats reached asymptotic performance in less than 50 trials after switching contingencies. Surprisingly, performance for extreme (non-reversing) frequencies on the initial trials of a block was higher than the asymptotic level. This can be explained by a category boundary that varies with changes in contingencies, modifying its distance to extreme frequencies as reversals are achieved.

To unveil the neuronal mechanisms implementing these rapid changes in behavior, we recorded the activity of single cells from auditory cortex and their striatal targets during changes in categorization contingencies. We found auditory cortical neurons whose response to a given sound was modulated by the sound-action association. Interestingly, several neurons showed increased responses to non-reversing sounds on the initial trials of a block, providing a neural correlate of improvements in performance during this period. Similarly, electrophysiological recordings from neurons in the auditory striatum displayed choice-related activity that was dependent on the stimulus identity. These observations suggest a role for cortico-striatal projections in flexibly establishing associations between sounds and actions.
Auditory object recognition using sparse spike codes and network dynamics

Dezhe Z. Jin and Phillip Schafer

Department of Physics, The Pennsylvania State University, University Park, PA 16802, USA, email: djin@phys.psu.edu

Understanding how the auditory system recognizes auditory objects such as speech and animal vocalizations is one of the central goals of systems neuroscience. Auditory signals are inherently dynamic. Thus, auditory object recognition requires integration over time, typically in the range of hundreds of milliseconds to a few seconds for speech and animal vocalizations. We propose a model that integrates spikes of sparsely firing feature detectors using network dynamics. A cochlea model is used to decompose the sound into the auditory nerve firing patterns. These patterns are transformed into sparse spikes of feature detectors, which are constructed using independent component analysis. The spatial-temporal spikes of the feature detectors are integrated using a network of neurons connected into a chain-like structure. These neurons receive ascending inputs from the feature detectors. The patterns of these feed-forward connections encode the specific auditory object to be recognized by the network neurons. The spiking activity of the network neurons, determined by the interaction between the sound-driven spikes of the feature detectors and the intrinsic dynamics of the network, signals the recognition or rejection of the sound signal as the encoded auditory object. The feed-forward connections can be learned through auditory experience.
Comparison between single neuron responses to cochlear implant and acoustic stimulation in auditory cortex of awake primate

Luke Johnson, Charley Della Santina, Xiaoqin Wang
Johns Hopkins University, Baltimore MD

A cochlear implant (CI) converts sounds to electric pulses that directly stimulate the auditory nerve via an electrode surgically implanted in the inner ear. Although cochlear implants have been successful at providing auditory sensation to many individuals with profound hearing loss, many further improvements are needed. Understanding how the brain processes cochlear electric stimulation compared to acoustic stimulation will guide future CI technology improvements.

We have developed a unique non-human primate CI model, the common marmoset (Callithrix jacchus). In our study, a multi-contact CI electrode was chronically implanted in one cochlea while the other cochlea remained acoustically intact. This enabled us to record and compare on a neuron-by-neuron basis responses to acoustic, electric or electric/acoustic combination stimuli. We recorded single unit responses in primary auditory cortex (A1) to electric current pulses and acoustic clicks delivered at repetition rates of 2-256 pulses per second (pps). We then examined each neuron’s responsiveness to CI electric stimuli compared to acoustic stimuli. Neurons responsive to both acoustic and electric pulses showed similar response profiles, in that neurons that showed synchronized responses to electric stimuli also showed synchronized responses to acoustic stimuli, while neurons with non-synchronized responses to CI stimuli generally showed non-synchronized responses to acoustic stimuli as well. These findings suggest that many A1 neurons utilize similar coding schemes to represent time-varying CI and acoustic stimulation.

However, important differences in A1 responsiveness to CI and acoustic stimuli were observed. We found that in general fewer neurons are activated by CI stimulation compared to acoustic stimulation and, intriguingly, the lack of responses to CI stimuli is more prominent for ipsi-lateral ear stimulation. This may be explained in part by broader cochlear excitation areas caused by electric stimulation compared to acoustic stimulation, and that many cortical neurons have narrow tuning and sideband inhibition. Therefore for a given neuron, the CI stimulation may either not excite the appropriate frequency-place area of the cochlea, or evoke too many inhibitory inputs that prevent a neuron from being
activated. This explanation is supported in part by the finding that neurons that are non-responsive to CI stimuli have narrower tone-tuning curves compared to neurons responsive to CI stimuli. These findings suggest that CI stimulation schemes that utilize more focused cochlear electric stimulation may be more effective in activating cortical neurons.
The segregation of simultaneous broadband sources in elevation using envelope cues: behavior and modeling

Jeffrey S. Johnson (1,2), Kevin N. O’Connor (1,2,3), Mitchell Sutter (1,2,3)

(1) University of California, Davis; (2) Center for Neuroscience; (3) Department of Neurobiology, Physiology, and Behavior

“Cocktail party” segregation refers to the ability of the auditory system to segregate multiple, simultaneous sound sources. Source azimuth can be determined using interaural time and intensity cues, but monaural pinna-based spectral notch cues are needed to determine source elevation. Cocktail party segregation of broadband sounds in elevation is particularly complicated because notch-based spectral features blend at the pinnae, and the ability to perform such a segregation has not previously been found. Here we show that humans are capable of segregating isospectral broadband noises in elevation using temporal envelope cues by asking them to identify the up/down direction of an amplitude-modulated noise when an unmodulated distracter is simultaneously presented from the opposite direction (speakers +/- 20 degrees, located on vertical midline). Some subjects perform at high accuracy for modulation frequencies up to 250 Hz, while others perform worse than chance, and a cluster analysis strongly suggests that these subjects comprise two separate groups. A model of the outputs of inferior colliculus type-O cells, monaural cells involved in spectral notch detection, suggests that the balance of their inhibitory and excitatory inputs could potentially explain both correct and incorrect response modes.
TITLE: HEARING-IMPAIRMENT IN ADULT FERRETS INDUCES PARTIAL CROSSMODAL CONVERSION OF CORE AUDITORY CORTEX.

AUTHORS: Les Keniston, Brian Allman* and M Alex Meredith,

VCU School of Medicine, Richmond, VA and

*Center for Hearing and Deafness, State University of New York at Buffalo.

ABSTRACT:

Most experimental studies of crossmodal plasticity have examined subjects that experienced complete sensory loss early in life. However, many adults suffer from partial sensory loss and few crossmodal effects have been identified in this population. Therefore, the present experiment was initiated to measure the extent of crossmodal reorganization that might occur in response to partial hearing loss in adult animals. The core auditory cortices were examined (multichannel single-unit recording) in adult (>200 days old), male ferrets that were ototoxically hearing impaired (n=5; avg. 49dB threshold) or hearing controls (n=5; <15dB threshold). Computer-controlled somatosensory, visual, and auditory stimuli were presented (50x each) alone and in combination while responses of neurons in primary auditory (A1) and anterior auditory field (AAF) were recorded (part deaf = 284 neurons; hearing = 311). Neurons responding only to acoustic stimulation were reduced in the hearing impaired animals (from 74% to 33.3%). In contrast, the proportion of neurons activated by more than one sensory modality increased >3 fold (from 13.3 to 55.5%), of which trimodal neurons represented the largest proportion (from 0% to 23.4%). These data confirm that crossmodal plasticity occurs within the auditory cortex of hearing-impaired adults (Allman et al., 2009) and establishes that even partial hearing loss increases crossmodal responsiveness in A1/AAF significantly. As a consequence, rehabilitative therapies in hearing-impaired adults will be substantially mitigated by the complicating factors of crossmodal (substitution of vision or touch) or multisensory (combination of vision or touch with remaining audition) plasticity.
Harmonic preference in the lateral belt of rhesus monkey auditory cortex

Y. KIKUCHI1,2, B. HORWITZ3, M. MISHKIN4, J. RAUSCHECKER2;

1Inst. of Neurosci., Newcastle Univ., Newcastle Upon Tyne, United Kingdom; 2Dept. Neurosci., Georgetown Univ., Washington, DC; 3NIDCD, Bethesda, MD; 4NIMH, Bethesda, MD

In a previous report (Kikuchi et al., 2009), we showed that neurons in the auditory belt areas appear to serve as high-resolution filters that extract specific harmonic features based on the relationship between the neuron’s peak frequency and fundamental frequency (F0). In the current study, we examined the single-unit activity and the local field potentials (LFP) recorded simultaneously in the core and lateral belt (LB) in the rhesus monkeys as they passively listened to pure tones and pitch-shifted monkey vocalizations (coos). The latter consisted of complex-tonal segments, in which F0 was matched to each pure-tone stimulus. Once we identified the neuron’s best frequency (BF) with pure tones, complex acoustic stimuli with the same center frequency but varying timbre were introduced by subtracting F0 and/or harmonics from the stimuli. In both animals, latencies to pure-tone stimuli at the BF were significantly longer in LB than in core (59 ms vs. 38 ms, respectively; P < 0.001). On the other hand, latencies at F0 did not differ between the two regions, and in LB, latencies to F0 (43 ms) were significantly shorter than those to BF (59 ms, P < 0.01). Furthermore, when measuring intervals between multiple peaks in the response to the tones, the proportion of harmonic intervals (perfect fifth, perfect fourth) was greater in LB than in core (50 % vs. 27 %; P < 0.01). At the population level, both the core and LB showed a burst of high-gamma (90-150Hz) and low-gamma (30-70 Hz) activity that were significantly greater to the harmonic stimuli compared to those to the pure tones (P < 0.01). Our results suggest that harmonic features, such as pitch and frequency interval of complex tones were formed already at relatively early stages of the auditory cortical pathway sharing the common synaptic input from the medial geniculate nucleus and provide preliminary evidence for the neural basis of harmonic preference in the LB.
Every day hearing activates large numbers of neurons in the central auditory system. It is not clear how many neurons are needed to generate an auditory percept or how activity among these neurons leads to useful behavior. The map plasticity generated by frequency discrimination training seems to suggest that many neurons contribute to the perception of even simple sounds. Recent observations that map plasticity aids discrimination learning but is not needed to perform the discrimination suggest that a new model of brain function is needed. In the Expansion-Renormalization Model large numbers of neurons are engaged during the early stages of learning so that a subpopulation of neurons that is most effective in performing the task can be identified. This model predicts that small numbers of neurons could generate phantom percepts, which might be reversed using large scale plasticity. Pairing brief pulses of vagus nerve stimulation (VNS) with tones is sufficient to eliminate both the neural and behavioral correlates of tinnitus in rats. Encouraging early results in tinnitus patients suggest that VNS-directed neural plasticity might be useful for treating other common neurological disorders, including stroke, dyslexia, and autism. Restoring normal function is likely to require a detailed understanding of the mechanisms relating neural activity to behavior. Similarities in animal and human speech sound processing abilities suggest that it will be possible to conduct extensive preclinical testing of novel human therapies for speech processing disorders.
Subcortical representation of sound is enhanced in bilinguals
Krizman J, Shook A, Marian V, Skoe E, Kraus N.

While bilingualism is known to impact cortical function, the effect of multiple language experience on subcortical processing is unknown. In the current study, we aim to investigate whether bilinguals and monolinguals process sounds differently at the subcortical level, and if these differences relate to cognitive abilities associated with executive function. To assess the impact of bilingualism on sound processing and cognitive abilities, auditory brainstem responses and behavioral measures of auditory processing and executive function were obtained in 60 adolescents of varying second language proficiency, matched on IQ and socioeconomic status.

Bilinguals, when listening to the syllable ‘da’ presented in six-talker background babble demonstrated more robust encoding of the stimulus compared to monolinguals. Subcortically, the spectral amplitude of the fundamental frequency (F0), a feature known to underlie pitch perception and grouping of auditory objects, was more robust in bilinguals compared to monolinguals. Behaviorally, bilinguals demonstrated perceptual advantages in the sensory processing of basic acoustic components, such as timing and frequency, which are fundamental to accurate speech perception. Specifically, bilinguals showed enhanced discrimination of simple, non-linguistic sounds as assessed by a measure of temporal resolution (backward masking) and a measure of frequency discrimination. Relative to monolinguals, cognitive advantages were seen in bilinguals in both the auditory and visual domain on a behavioral measure of executive function (i.e. sustained, focused attention). Performance on both the visual and auditory measures of executive function was correlated with subcortical F0 amplitude. The brainstem response to complex sounds was more robust in bilinguals than monolinguals. This enhanced neural encoding related to better performance on tests of executive function and is consistent with the notion that enhanced executive function may drive the strength of subcortical stimulus representation. The present findings suggest an enhancement for bilinguals in the neural processing of specific sound elements that relate to auditory perception and cognitive abilities. The subcortical strength of F0 encoding is known to be influenced by sensory and cognitive processes. An interaction among cognitive, sensory, and subcortical functions may provide bilinguals an advantage in listening situations where enhanced auditory acuity and cognitive control are essential to communication.
Dissociating representation of spectrotemporal features from perceived unpleasantness of aversive sound stimuli in the amygdala: a human fMRI study

Sukhbinder Kumar¹,², Katharina von Kriegstein³, Timothy D Griffiths¹,²

¹ Institute of Neuroscience, Medical School, Newcastle University, Newcastle upon Tyne, NE2 4HH (UK)
² Functional Imaging Laboratory, Wellcome Centre for Neuroimaging, 12 Queen Square, London, WC1N 3BG (UK)
³ Max Planck Institute for Human Cognitive and Brain Sciences Stephanstraße 1A, 04103 Leipzig, Germany

Abstract

Some sounds (e.g. fingernails scratched on blackboard) are perceived as highly unpleasant. In a previous study (Kumar et al, 2008) we showed the unpleasantness of these sounds can be predicted from their spectro-temporal features. In particular, it was shown that frequencies in the band (2500 – 5500 Hz) that are modulated with low temporal modulations (1- 16 Hz) correlated with the subjective rating of unpleasantness. A number of studies (Zald and Pardo, 2002; Mirz et al, 2000) in humans have previously shown activation of the amygdala in response to unpleasant sounds. Brain response to an aversive sound stimulus, involves a number of processes at least two of which, among many others, include (i) representation of the spectro-temporal features of the aversive sound stimuli and (ii) the subjective experience of unpleasantness. Previous studies using sounds as aversive stimuli have not made distinction between these two dimensions of the response. It is, therefore, not clear whether spectro-temporal features of stimuli are encoded in the amygdala or it mediates only the subjective experience of unpleasantness or it is involved in both of these processes. In order to investigate these questions, we used functional magnetic resonance imaging (fMRI) to measure brain activations in response to a range of sounds which are rated by subjects from least unpleasant to highly unpleasant. In addition, we also extracted spectro-temporal features of the same sounds to determine if activity of the amygdala is modulated by these features. Our results show that (i) the spectro-temporal features of the sounds are encoded in the amygdala bilaterally (ii) activity in the right amygdala only correlates with the
perceived unpleasantness of sounds (iii) activity of the auditory cortex is modulated both by the perceived unpleasantness of sounds and the spectro-temporal features of the sounds that are relevant for perceived unpleasantness.

References


Kusmierek P, Ortiz M, Rauschecker J P

Population analysis reveals specialization for “what” processing in early rostral areas of macaque auditory cortex.

Early work on lateral belt of macaque auditory cortex has reported a relative preponderance of neurons selective to communication calls in anterior belt (area AL) compared to caudal belt (CL; Tian et al 2001). Together with results from tracer studies and from human imaging, this report has contributed to the suggestion of an anteriorly directed “what” stream for sound identification. However, the original finding was challenged by a study that reported no difference in call selectivity between CL, middle lateral belt (ML), and core areas A1 and R (Recanzone 2008).

We analyzed neural population responses to artificial sounds (tones [PT], ⅓-oct and 1-oct band-pass noise bursts [BPN]), and to natural sounds (monkey calls [MC], environmental sounds [ES]) in regions R (areas R+RM), M (A1+MM) and C (CM) of four rhesus monkeys. The stimuli were 151 ms or longer, thus, all analyses covered 160 ms after stimulus onset. Similarity matrices and hierarchical clustering based on distances between population responses to stimuli showed that PT/BPN center frequency (CF) strongly affected population responses, and that population responses were more similar within each natural sound class than between natural classes, or between natural and artificial sounds.

Using k-means clustering, the stimuli were grouped into k clusters by minimizing distances (calculated from population responses) to cluster centers, with k=4 (ES, MC, low-CF PT/BPN, high-CF PT/BPN) or k=5 (mid-CF PT/BPN added). Percent correct classification (PC) reached 57-90% compared to 35-40% PC based on pretrial. Then, k-means clustering was done separately for each region in eight 20-ms segments. PC based on region C (C-PC) increased quickly from the pretrial level, followed by M-PC and then R-PC, confirming latency differences between the regions. From 60-80 ms after stimulus onset, R-PC was significantly higher (76-95% for k=4) than PC obtained from other areas (M-PC: 58-74%). The advantage of R-PC held even when only responses to MC and ES were analyzed with k=2. Clustering of stimuli based on acoustic measures approached but did not reach the best R-PC level. Patterns seen in similarity matrices based on neural responses were never fully replicated in similarity matrices based on acoustics.

Thus, anterior specialization for “what” processing can be detected at early stages of cortical processing. The rostral regions do surpass more posterior regions in the ability to identify/classify sounds. The relatively slow onset of high R-PC and poorer classification based on acoustics suggest that feedback connections from areas further down the “what” stream may contribute to stimulus classification at the level of R.

Supported by: NIH Grant R01-NS052494 to JPR
The auditory cortex plays an important role in detecting behaviorally relevant stimuli, such as species-specific vocalizations. Previous studies have demonstrated in a mouse model of maternal behavior that mothers undergo long-term cortical changes in the way pup vocalizations are represented (Liu et al, 2006; Liu and Schreiner, 2007; Galindo-Leon et al., 2009). Mouse pups emit ultrasonic vocalizations when isolated from the nest, and unlike pup-naïve virgins, mothers and virgins with pup-caring experience (cocarers) recognize these calls as behaviorally relevant (Ehret and Haack, 1982; Ehret et al. 1987). However, the degree to which experience and maternal hormones interact to produce this sensory plasticity is not well understood.

Recording from the auditory brainstem and cortex of mothers post-weaning (21 days of pup experience), we found that mothers had significantly earlier subcortical sound-evoked responses and stronger auditory cortical pup call-evoked inhibition compared to pup-naïve virgins. The shift in auditory brainstem responses was not driven solely by an effect of maternal levels of estrogen, as revealed by data collected from estrogen- and blank-implanted, ovariectomized virgins. Instead, data from post-weaning cocarers suggests that pup experience plays a strong role in this shift, since their responses were closer to mothers. Furthermore, the cortical differences between post-weaning mothers and pup-naïve virgins likely involves maternal hormonal mechanisms in addition to pup experience, since data from post-weaning cocarers was intermediate between mothers and pup-naïve virgins. Our results therefore provide further evidence that the maternal experience imparts lasting changes in auditory processing of infant cues, both cortically and subcortically, with roles for both pup care and maternal hormones in the maintenance of these changes.

Acknowledgements: Funding provided by NIH grant DC008343, and the Center for Behavioral Neuroscience under the STC Program of the National Science Foundation (IBN-9876754), and the NIH Institutional Research and Academic Career Development Award K12 GM000680.
Neural Deficits in Auditory Temporal Processing in Auditory Thalamus of Ectopic BXSB/MpJ Mice

Jennifer F. Linden1,2 and Lucy A. Anderson1,
1 Ear Institute and
2 Dept. Neuroscience, Physiology and Pharmacology, University College London, UK

How does the mammalian auditory system perform the complex temporal processing required for the perception of sound? The answer to this question may become clearer if we can understand how malfunctions of the auditory system cause auditory temporal processing deficits. In humans, auditory processing deficits such as difficulty perceiving rapidly changing sounds are often associated with developmental disorders, including some forms of dyslexia. Inbred mice of the BXSB/MpJ strain are one of the best animal models of auditory temporal processing disorder. Approximately half the animals within a BXSB/MpJ litter develop malformations in their prefrontal and parietal cortices whereby nests of neurons (ectopias) form in layer I. These ectopias are similar to those observed in some humans with developmental disorders.

Previous behavioural work has shown that despite having normal hearing sensitivity, ectopic male BXSB/MpJ mice have auditory temporal processing deficits, such as difficulty perceiving rapid changes in sounds. In other animal models of auditory temporal processing disorder, neocortical micro-abnormalities like ectopias have been associated with anatomical abnormalities in the auditory thalamus. Here, we report evidence for physiological abnormalities in the auditory thalamus of ectopic male BXSB/MpJ mice. We recorded from single neurons and multi-neuron clusters throughout the auditory thalamus of male BXSB/MpJ mice, and also measured auditory brainstem responses (ABRs). Although ABR thresholds, wave amplitudes, and wave latencies were normal in ectopic animals, we found neural deficits in auditory temporal processing in the thalamus. In particular, neurons recorded from ectopic males had significantly attenuated responses to brief gaps in noise, compared to neurons recorded from non-ectopic animals. However, responses to other types of temporally varying stimuli (e.g. click trains) were normal, as were other response properties such as first-spike latency distributions and response thresholds. Thus, ectopic male BXSB/MpJ mice appear to have a specific central auditory deficit in temporal processing, which is evident at the level of the thalamus.

Supported by: Wellcome Trust 084364/Z/07/Z
Psychophysical behavior during an auditory frequency contour discrimination task
Andrew Liu, Joji Tsunada, Josh Gold, Yale Cohen

Submitted by: Andrew Liu, Bioengineering
Advisor: Dr. Yale Cohen

Perceptual decision-making is a deliberative process that converts incoming sensory information into a categorical judgment. This process is often modeled as the sequential sampling of sensory evidence over time, which can account for the trade-off between speed and accuracy that is inherent to many (but not all) visual tasks. In contrast, little is known about the temporal dynamics of auditory decision-making. Because auditory processing is inherently temporal, specializations in the auditory system are likely to have substantial effects on these dynamics. In this study we tested how changes in the spectrotemporal features of an auditory stimulus, which can dramatically affect its perceptual qualities, can also affect the speed and accuracy of human subjects making categorical judgments about that stimulus.

We designed a novel auditory-decision task that required subjects to report whether a sequence of tone bursts was increasing or decreasing in frequency. We systematically varied two features of these tone-burst sequences. First, we varied stimulus coherence, which affected the difficulty of the perceptual judgment. The coherence determined the probability that each tone burst in a given sequence corresponded to a fixed frequency offset (increment or decrement) from the previous burst. When the coherence was 100%, all bursts were offset in this manner. When the coherence was 0%, no bursts were offset in this manner and instead were chosen at random from a pre-defined set of frequencies. Second, we varied the time between sequential tone bursts (the inter-burst interval). This manipulation affected the perceptual “smoothness” of the frequency contour, making it sound like one smoothly varying sound for short intervals, or multiple, sequential bursts for longer intervals.

We found that, as coherence increased, decision accuracy increased and mean reaction time decreased. These results were well described by a model in which sensory evidence is accumulated over time to a fixed boundary, consistent with certain visual decisions. Moreover, as the inter-burst interval increased (i.e., the perceptual smoothness decreased), accuracy decreased and reaction time increased (when considering overall stimulus time) or decreased (when considering only the total number of bursts and not the inter-burst interval) for a given coherence. Preliminary analyses suggest that these reaction-time differences reflected changes in the bound height in the model, not the rate of evidence accumulation. Thus, spectrotemporal regularities of an auditory stimulus that can affect its perceptual qualities can also affect the process of forming decisions based on that stimulus.
INCREASING SPECIFICITY FOR COMPLEX ACOUSTIC STIMULI TOWARDS THE TEMPORAL POLE OF THE CAT CEREBRUM

Stephen G. Lomber and Amee J. McMillan
Cerebral Systems Laboratory, Centre for Brain and Mind,
The University of Western Ontario, London, Ontario, Canada.

Electrophysiological, behavioural, and connectional studies have identified a ventral or “what” processing stream within extrastriate visual cortex of the human, monkey, and cat. This pathway arises in occipital cortex and projects anteroventrally across the temporal lobe. Furthermore, studies have shown that visual areas along this pathway become responsive to increasingly complex visual stimuli. Similarly, in auditory cortex, electrophysiological studies suggest that areas beyond primary auditory cortex (A1) become responsive to increasingly complex acoustic stimuli along the sylvian gyrus of the feline temporal lobe. To examine this idea, we tested the hypothesis that ventral, but not dorsal, areas of the cat’s temporal lobe have greater specificity for complex acoustic stimuli. We trained four mature cats to discriminate sounds in a two-alternative forced-choice apparatus. The animals concurrently learned to discriminate three classes of sounds: tones, narrow-band bursts, and conspecific vocalizations (“meows”). With criterion performance of 70% correct on three consecutive days, we identified that conspecific vocalizations were learned the fastest, while tonal discriminations often required twice as much time to master. After training, cooling loops were bilaterally placed over primary auditory cortex (A1), second auditory cortex (A2), temporal cortex (area T), and insular cortex (area IN) to permit their temporary and reversible deactivation. The animals were then tested while each of the areas was bilaterally or unilaterally deactivated. Presentation of the three classes of stimuli was randomly presented within each testing session. Bilateral deactivation of A1 resulted in discrimination deficits on all three stimulus classes. Bilateral deactivation of A2 caused deficits only for the narrow-band burst and conspecific vocalization classes. Bilateral deactivation of area T resulted in deficits restricted to the conspecific vocalizations. Unilateral deactivation of left, but not right, area T caused deficits during conspecific vocalization discriminations. Therefore, the present investigation provides evidence for the lateralization of conspecific vocalization discrimination in the left hemisphere. Bilateral deactivation of area IN did not produce deficits on any of the three stimulus classes tested. Overall, specificity for increasingly complex acoustic stimuli was identified along the temporal lobe of the cat. The results of this study indicate a “what” processing pathway in auditory cortex of the cat that arises in primary auditory areas and radiates down the temporal lobe through A2 and into area T.

This work was supported by grants from the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada, and the Canada Foundation for Innovation.

Registration Information:

Steve Lomber
Centre for Brain and Mind
Department of Physiology and Pharmacology
The University of Western Ontario
Medical Sciences Building, Room 216
London Ontario N6A 5C1
Canada
519-661-2111 x24110
steve.lomber@uwo.ca
Discrimination of acoustic sequences in songbird auditory forebrain.
Lu, Ziv and Vicario

Perception of language and music depends on discriminating the order of acoustic events at different time scales, but the neural mechanisms of sequence processing are not well understood. Songbirds provide an excellent potential model because their complex learned songs consist of multiple notes and syllables in patterned sequences that resemble human speech. These signals must be discriminated because they communicate individual identity and other variables that are important for social and reproductive behavior that contributes to the survival of the species. The current study explores the neural processing of sound sequences in two auditory areas in the avian forebrain, the caudo-medial nidopallium (NCM) and caudal mesopallium (CM), that have been shown to encode long-term memories for individual songs. Multi-unit recordings were made from 16 electrodes placed in NCM and CM of awake, restrained adult male zebra finches (N=5). The birds heard 20 repetitions of an acoustic sequence consisting of four artificial song notes to induce stimulus-specific adaptation, a reduction in multi-unit neural responses that occurs in these auditory areas to familiar sounds. When the same notes were presented in a new sequence, responses increased in NCM and to a lesser degree in CM, indicating that the changed sequence was detected by these neurons. We conclude that neural responses in NCM and CM are not only sensitive to the acoustic features of individual notes and syllables, but also to the temporal structure of sounds over longer time scales (up to hundreds of milliseconds).
Factors accounting for variation in the degree of contralateral preference in human auditory cortical processing of binaural cues: A functional magnetic resonance imaging study

Prior human neuroimaging studies have yielded incongruent findings regarding the degree of contralateral preference in human auditory cortex for the processing of interaural time differences (ITD) and interaural level differences (ILD), binaural cues important for the localization of sound. Factors that might influence the variability among those results include the magnitude of binaural cues and the possible presence of stimulus-specific adaptation. Specifically, some data suggest that contralateral processing dominance may be maximal at middle binaural cue values (e.g., ITDs of 500-700 µs, ILDs of 10-20 dB) and minimal at extremely large or small values (e.g., ITDs < 1000 µs, ILDs < 5 dB or > 25 dB). The reduced contralateral preference at large binaural cue values could be explained for ITD by the increased ambiguity of the cue at large values [see von Kriegstein K et al. J Neurophys 100:2712-18, 2008], but the same effect would not be expected for ILD. Preliminary findings from our lab, however, do show a similar result for ILD, in which “U-shaped” ILD response-functions are seen, with increased bilateral activation at 30 dB.

This experiment, therefore, employed functional magnetic resonance imaging (fMRI) to measure human AC responses to variation in ILD values, using a continuous imaging paradigm (TR=2s, 42 slices, 2.75 x 2.75 x 3mm resolution, 3T). Stimuli consisted of narrowband click trains (carrier frequency of 4000 Hz, half-maximal bandwidth of ~1.8 kHz, and interclick interval of 2 ms) varying across ILD (0, ±5, ±10, ±15, ±20, and ±30 dB). Each stimulus was presented for 1-s, with the interstimulus interval jittered from 1 to 5-s, in an event-related paradigm that employed a continuous carryover design [Aguirre GK Neuroimage 35:1480-94, 2007] to counterbalance stimulus presentation in order to assess potential effects of stimulus history on neural response (i.e., stimulus-specific adaptation). Average binaural listening level was held constant at 80 dB SPL. Participants were asked to visually fixate on a center cross in order to minimize the effect of eye movements, and to indicate by a right-handed button press infrequent 10% pitch changes unrelated to the experimental question. It is expected that contralateral dominance in the AC for auditory spatial processing will be a subtle effect, as suggested by neurophysiological recordings in non-human primates [Werner-Reiss U and Groh JM J Neurosci 28: 3747-58. 2008], which may depend critically on hemisphere, size of ILD, degree of stimulus-specific adaptation, and regional variation across AC fields. Work supported by NIH R03-DC009482-02S1.
Invariant cortical representation of attended speaker in multitalker speech perception

Nima Mesgarani & Edward Chang

1. Keck Center for Integrative Neuroscience
2. Departments of Neurological Surgery and Physiology

A unique and defining property of human speech perception is the ability to robustly process speech sounds in the context of noisy and interference-filled acoustic conditions. A common, everyday condition is the multi-speaker environment where selective listening is required for listening, also known as the "cocktail party effect". The mechanism by which the human auditory system carries out sound processing under these conditions is largely unknown. An attractive mechanism for speech encoding in multi-speaker environments is the implementation of dynamic "top-down" modulation of attention towards the intended signal. In this study, we used high-resolution intracranial direct recordings (electrocorticography) from the superior temporal gyrus (STG) in patients with intractable epilepsy in order to investigate the neural correlates of auditory selective attention. The behavioral paradigm was based upon the Coordinate Response Measure (CRM) corpus which is widely used for multi-speaker communications research. The patients are instructed to report the color and number associated with a call sign (e.g. "Tiger") in a mixture of two speakers, without knowing a priori which speaker will be the target in a given trial. We used stimulus reconstruction method from the high-gamma envelope of neural responses to investigate the encoding of the spectrotemporal features of attended speech. The reconstructed spectrograms from the same acoustic sound mixture, but in two different attention conditions resembled the spectrogram of the target speaker in isolation, indicating an enhanced neural representation of the attended voice. In addition, using a linear classifier trained on the representation of single speakers, we successfully decoded the spoken words and the identity of the target speaker from the responses to the mixture. Finally, we investigated how the tuning properties of different areas in STG changed rapidly in order to enhance the features of the attended voice while at the same time suppressing the representation of the distractor speaker. Our results show that attention modulates the neural representation of speech in STG in order to provide an invariant and robust representation of the attended speaker.
Context-specific responses in primate prefrontal cortex neurons during natural behavior: antiphonal calling

Cory T. Miller, A. Wren Thomas & Ryan Morrill
UCSD

Primates evolved sophisticated behaviors in order to effectively navigate their respective social and ecological landscapes. While these behaviors clearly had a significant impact on the evolution of neural mechanisms in primate neocortex, relatively little remains known about how the brain functions during natural behaviors. Amongst the most significant of these behaviors are those involved in communicating with conspecifics. We previously found that neural responses in primate frontal cortex during natural vocal interactions were behaviorally dependent. We found, for example, that the strength of the neural response was positively correlated with the number of consecutive calls in an antiphonal calling bout. Here we extend this work to investigate the various contextual differences that may underlie the observed neural response during this natural behavior. We recorded the single-unit activity of frontal cortex neurons in common marmoset monkeys (*Callithrix jacchus*) in the following contexts. First, antiphonal calling: This natural, species-typical vocal behavior involves the reciprocal exchange of vocalizations between conspecifics. Producing an antiphonal call is dependent upon first hearing a particular vocalization, a phee call, and producing the same call type in response. Second, head fixed stimulus presentation: Here we presented a series of vocalization and white noise stimuli while animals were restrained in a primate chair. Third, freely-moving stimulus presentation: In this condition, we presented subjects with a series of vocalization and white noise stimuli while subjects were freely moving. The same stimulus sets were used in all three of these behavioral contexts and only neurons that were recorded in all three contexts are used in the analysis. Analyses focus on sensory stimuli across each of these contexts to determine the effect of stimulus presentation (repetitive vs sparse), mobility and communicative involvement on neural responses.

This work supported by grants from the NIH R00 DC009007
Tonotopic map of human inferior colliculus unraveled by functional MRI at 7T

Michelle Moerel¹,², Federico De Martino¹,²,³, Kamil Ugurbil³, Pierre-Francois van de Moortele³, Elia Formisano¹,², and Essa Yacoub³

¹Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, ²Maastricht Brain Imaging Center (MBIC), Maastricht, The Netherlands, and ³Center for Magnetic Resonance Research (CMRR), University of Minnesota Medical School, Minneapolis, Minnesota.

Tonotopy - the orderly spatial arrangement of neural tuning to sound frequency - is a general organizational principle of the auditory system that arises in the cochlea and is preserved throughout various auditory relays [1]. In humans, the spatial layout of tonotopic cortical maps has been mapped using functional MRI (fMRI) [2]. However, the limited spatial resolution and brain coverage of previous investigations has excluded the possibility of examining tonotopic maps in the human midbrain.

We used high field fMRI (7T) to investigate tonotopy in human inferior colliculus (IC), acquiring two sets of high resolution (voxel dimension = 1.5 mm isotropic) data while subjects (n = 3) listened to tones in three frequency ranges (exp. 1; 0.5 KHz/ 1.5 KHz/ 2.5 KHz; block design) or recordings of 144 natural sounds (exp. 2; e.g. speech, music; fast-event related design). Maps were obtained by colour-coding the frequency to which voxels responded strongest (exp.1) [2] or through a novel modelling approach linking fMRI data to the frequency spectrum of natural sounds (exp.2).

We observed a low-to-high tonotopic gradient in dorsolateral to ventromedial direction (with respect to brainstem), consistent between experiments (Fig. 1, sagittal section for subject 1, for exp.1 [a] and exp.2 [b]) and subjects (group maps in Fig. 1, for exp.1 [c] and exp.2 [d]). This indicates that - in human IC - sound frequency is represented topographically. The gradients orientation corresponds with observations in the IC of several animal species [1,3], indicating that the tonotopic spatial arrangement is preserved in the mammal auditory midbrain.

Tonotopy mapping in human IC is a fine example of possibilities enabled by high field fMRI. Its combination with natural sounds and advanced modeling techniques can be used for examining other processing principles of the human auditory midbrain non-invasively and at high spatial resolution.
The Production and Perception of Ultrasonic Vocalizations in CBA/CaJ Mice

Erikson G. Neilans, Kelly E. Radziwon, and Micheal L. Dent

Department of Psychology, University at Buffalo, The State University of New York, Buffalo, NY 14260

Previous studies have found that both male and female mice produce ultrasonic vocalizations (USVs) under many circumstances, yet the function of these USVs remains unclear. Understanding USVs in mice is beneficial to understanding the genetic contributions underlying the production and perception of auditory communication. In this study, we examined whether vocal cues alone could elicit vocalizations in CBA/CaJ mice, and whether the mice could discriminate between calls produced by conspecifics. In experiment 1, a variety of USVs recorded from both males and females were played back to different male and female mice, whose subsequent vocal responses were recorded. Both male and female mice readily called back to playback USVs, and female mice vocalized as much as males when presented with conspecific calls.

In experiment 1, it appeared that mice emitted several types of USVs that differed in spectral and temporal qualities. In experiment 2, we tested the discrimination ability of new mice to calls that we recorded in experiment 1. The mice were trained, using operant conditioning procedures and positive reinforcement, to discriminate target calls from repeating background calls. Certain calls were difficult to discriminate between, specifically target calls that fell within the same spectral range as the repeating background. Additionally, temporally reversed, gapped, doubled-duration, and synthetic calls were difficult to discriminate from unaltered calls, suggesting again that spectral range is important to mice in the perception of USVs. These experiments highlight the importance of using the animals to categorize their own calls instead of human experimenters.
Mamiko Niwa

Activity correlated to animals' decision in primary auditory cortex (A1)

We found A1 activity correlated to animals’ decisions during amplitude modulation (AM) discrimination, possibly reflecting somatosensory, motor and/or attentional effects. Choice probability (CP) analysis was used to quantify the relationship between multiple- (MU) and single- (SU) unit spikes and the animals’ trial-by-trial decisions.

During the test stimulus, 20% of MUs and 14% of SUs significantly increased firing rate on trials where the animal reported AM compared to trials where they did not. We also found smaller, significant CP during the interval prior to the test stimulus, but after a standard sound, suggesting that the animals’ attention and/or state prior to the test stimulus may be a source of significant CP. During-stimulus CP significantly correlates with neurons’ AM sensitivity, which can be interpreted as attentional selection of cells that are better at AM discrimination.

We also found that activity after test stimulus offset was correlated with the animals’ behavioral response. Prior to the lever-release responses (‘detect-release’), 47% of MUs and 23% of SUs significantly increased firing rate on trials where the animal reported AM compared to trials where they did not. The majority of units that significantly increased firing rate before detect-release also increased firing before lever-release in a non-detection context. The magnitude of this CP also significantly correlated with neurons’ AM sensitivity, suggesting that somatosensory/motor-related activity is stronger in cells carrying the most task-relevant information. This could reflect a learned association between the stimulus and the auditory task, and shows the task-dependency of non-auditory influence on A1 neurons.
Regions of Human Auditory Cortex Exhibit Pitch-Selective Responses Across a Wide Variety of Sounds

Samuel Norman-Haignere (1), Josh McDermott (2), Evelina Fedorenko (1), Nancy Kanwisher (1)

(1) McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, USA, (2) Center for Neural Science, New York University, New York, USA

Pitch is a salient perceptual feature of many real-world sounds, including music and speech. Despite longstanding interest, evidence for a pitch center in auditory cortex has remained controversial. Some groups have reported elevated responses to sounds with pitch in a localized region of Heschl's gyrus, while others have reported anatomically distributed responses throughout auditory cortex. However, most studies have tested responses to a relatively small set of synthetic pitch stimuli, and it remains unclear how specific these prior results are to the stimuli used. Here we provide a more comprehensive test of neural pitch selectivity by examining fMRI responses to a broad range of natural and synthetic sounds. Six subjects were scanned while listening to stimuli drawn from 15 natural (e.g. orchestral music, speech) and synthetic (e.g. harmonic complexes, frequency-matched noise) sound categories that varied widely in pitch strength as well as other low- and high-level properties. In each subject tested, we observed two regions with consistent functional and anatomical profiles. A region in anterior auditory cortex, extending from lateral Heschl's gyrus into the planum polare, responded more strongly to sounds with pitch than to those without. This pitch preference was observed across many categories of natural and synthetic sounds, and the region's response to individual stimuli was well predicted by a simple measure of periodicity. A second region in the upper bank of the middle superior temporal sulcus was similarly selective for pitched sounds, but also showed an additional preference for speech that could not be explained by pitch (as measured by waveform periodicity). These results indicate that a small number of cortical regions respond preferentially to pitch across a wide range of natural and synthetic sounds.
Tonotopic organization of the human lateral superior temporal gyrus: Basic response patterns

Kirill V. Nourski¹, Mitchell Steinschneider², Hiroyuki Oya³, Hiroto Kawasaki¹, Matthew A. Howard III¹

¹ Department of Neurosurgery, The University of Iowa, Iowa City 52242, IA, USA
² Department of Neurology, Albert Einstein College of Medicine, New York 10461, NY, USA

Auditory cortex on the human posterolateral superior temporal gyrus (PLST) carries out sound processing within an interconnected system of core, belt and parabelt fields. The place of PLST within this hierarchy and its functional organization are not clear. Orderly distributions of best frequencies across neuronal assemblies provide a criterion for field differentiation within auditory cortex (Merzenich & Brugge, Brain Res. 1973, 50:275-96). Understanding basic response patterns elicited by pure-tone stimuli may aid delineation of PLST on physiological grounds.

Subjects were neurosurgical patients undergoing chronic invasive monitoring for refractory epilepsy. Stimuli were tone bursts ranging in frequency from 250 to 8000 Hz, and were delivered diotically via insert earphones during passive-listening experiments. Recordings were made over perisylvian cortex using high-density subdural grids. Responses were characterized by measuring high gamma (75-150 Hz) event-related band power (ERBP). Activation patterns across the recording grid in each subject were analyzed using a sparse logistic regression classification algorithm.

Pure-tone stimuli elicited robust responses in PLST, centered around the transverse temporal sulcus. Responses from individual recording sites exhibited frequency selectivity and were typically broadly tuned. Different frequencies elicited distinct activation patterns across PLST, with higher frequencies producing larger responses in more posterior sites compared to low-frequency tones. Often there was more than one gradient, suggesting multiple fields. Classification analysis of whole-grid ERBP data yielded above-chance performance, with maximum accuracy typically achieved within 100-150 ms after stimulus onset. Classification performance was statistically more robust in data recorded from the right hemisphere compared to left-hemisphere cases.

Results suggest that PLST contains multiple topographically-organized regions, possibly indicative of multiple fields. Broad frequency tuning indicates spectrally diverse, convergent inputs. Pattern recognition analysis indicates that PLST contains sufficient information for accurate classification of pure-tone stimuli. Higher classification accuracy in the right hemisphere is consistent with the concept of functional asymmetry of human auditory cortex.

Ongoing work (see companion abstract by Steinschneider et al.) is examining whether tonotopicity in PLST is intensity-invariant and whether it can predict encoding of more complex stimuli.

Supported by NIH RO1-DC004290, UL1RR024979, and the Hoover Fund.
Spectral contrast sensitivity of primary auditory cortical neurons: effects of bandwidth and ripple frequency. K. N. O’Connor1,2, P. Yin1,3, C. I. Petkov1,4 and M. L. Sutter1,2


It is well established that listeners’ ability to perceive communication sounds, such as speech, is enhanced as signal bandwidth increases (e.g., Noordhoek et. al. 1999; Hall et. al. 2008), though little is known of the neural processes underlying this improvement. We explored a possible cortical contribution by examining the sensitivity of primary auditory cortical (A1) neurons to spectral contrast (depth of modulation) in stimuli that could varied parametrically in bandwidth (BW) as well as spectral pattern, that is, ripple (sine-profile) frequency (SPF).

We recorded A1 single-unit (SU) potentials from awake, passively listening macaques. The stimuli were sinusoidal variations in power as a function of log frequency, spectrally windowed by a log-normal function (“spectral Gabors”) (O’Connor et. al. 2010). The signals (200-ms duration; 5-ms ramps) comprised 1024 pure tones at equal log intervals over 4 octaves, with randomized phases. The spectrum was centered over each neuron’s best frequency and its response was measured as a function of sinusoidal modulation depth (0-100%), at one (or more) combination(s) of SPF (0.75-17.25 cyc/oct) and BW (Gaussian SD = 0.25-8 oct). The effect of modulation depth on response rate was evaluated statistically using a bootstrapped, repeated-measures ANOVA. We also examined neurons’ spectral contrast sensitivity using a signal detection analysis. The probability distributions of trial-by-trial response rates to unmodulated and modulated stimuli were used to determine a sensitivity measure, the area under the receiver operating characteristic (ROC), at each depth (>0% modulation). These ROC values were then used to construct neurometric functions for each neuron, and detection thresholds (ROC area = 0.75) were estimated from these functions.

Of the 130 neurons effectively driven by spectral Gabor stimulation, 113 (87%) were tested on the spectral contrast experiments. Of these cells, 63 (56%) showed a significant effect of depth of modulation on spike rate (F ratio, P < 0.05). The likeliest outcome was an increase in rate to increasing spectral contrast, an effect displayed by 38 (60%) of the neurons showing significant effects. A few neurons showed a declining rate (4, 6.3%), or a non-monotonic response function (7, 11.1%). Neural sensitivity to spectral contrast (measured by the ROCs) was far better at larger BWs and lower SPFs. Neuronal thresholds were distributed above and below previously measured macaque behavioral thresholds for Gabor spectral contrast, suggesting that A1 population coding underlies spectral contrast sensitivity.
Interaural level difference (ILD) is an important cue for the localization of a sound source. The ILD information is first processed in the superior olivary nuclei in the brainstem and then conveyed to the IC in the midbrain. In the IC, ILD information is integrated and further processed through complex neural circuits. Synaptic processing or integration is critical for the information processing in the IC. However, there is little information on the discrete excitatory and inhibitory conductances generated by the inputs to IC neurons responding to binaural sound stimuli. Here, we studied the synaptic responses of IC neurons to ILD stimuli in mice with in vivo whole-cell recordings under voltage clamp.

We recorded the excitatory postsynaptic currents (EPSC) to best frequency (BF) tones from 43 IC cells. For 27 out of 43 cells, we also recorded the inhibitory postsynaptic currents (IPSC). All the cells had the EPSCs and IPSCs to monaural stimuli in the contralateral ear. Surprisingly, the 79.0% cells (34/43) also had the EPSCs to monaural stimuli in the ipsilateral ear, even though extracellular recordings suggest that many IC cells are contralaterally excited and ipsilaterally inhibited (EI cells). Among the cells that we successfully recorded the IPSC, 74% cells had IPSCs to monaural ipsilateral stimuli. In most cells, the EPSCs and IPSCs to ipsilateral stimuli were smaller than those to contralateral stimuli.

In 27 cells, we investigated synaptic responses to ILD stimuli where the average binaural level was constant. Most IC neurons showed the strongest synaptic currents at ILDs favoring the contralateral side. Interestingly, the EPSCs and IPSCs of these neurons were well balanced and their response patterns to ILD changes were highly correlated. In contrast to the majority, in other cells, the EPSC was largest around 0 ILD, but the IPSC was larger when the level was higher in the contralateral ear. In those neurons, EPSCs and IPSCs were poorly correlated.

These results showed that the binaural synaptic processes in the IC are far more complicated than expected from the extracellular responses. Most IC neurons did not have the simple synaptic input patterns predicted by a simple EI cell model. Instead, they had both EPSCs and IPSCs to contralateral and ipsilateral sound stimuli. This suggests that the balance and imbalance of the excitatory and inhibitory conductances, as well as the filtering by spike threshold, are important in shaping the ILD sensitivities of IC neurons.

NIH grant R01-DC000189 (DLO)
Cochleotopic mapping of macaque auditory cortex with functional magnetic resonance imaging at 3 Tesla

AUTHOR BLOCK: *M. ORTIZ¹, D. A. ARTCHAKOV¹, I. DEWITT¹, P. KUSMIEREK¹, D. CUI¹, J. VANMETER², J. P. RAUSCHECKER¹;
¹Dept. of Neurosci., ²Ctr. for Functional and Mol. Imaging, Georgetown Univ. Med. Ctr., WASHINGTON, DC

Abstract:
Macaque auditory cortex is organized into concentrically organized regions (core, belt, and parabelt), which can be distinguished neuroanatomically and electrophysiologically (Kaas & Hackett, 2000; Rauschecker et al., 1995). Microelectrode mapping further subdivides each region into cochleotopically organized fields, which are delimited from each other by reversals in best-frequency gradients along the caudal-to-rostral axis of the superior temporal gyrus (STG). Using pure-tone and band-pass noise stimuli under passive listening conditions, functional magnetic resonance imaging (fMRI) studies of monkey auditory cortex have found a similar organization (Petkov et al., 2006; Baumann et al., 2010; Tanji et al., 2010). In the present study, we measured blood oxygenation level-dependent (BOLD) responses to groups of low (0.5-1 kHz) and high (8-16 kHz) frequency band-pass noise (BPN) with 10-Hz sinusoidal amplitude modulation under an active discrimination task. Monkeys were trained to lie in a horizontal primate chair in prone (“sphinx”) position and fixated on a spot of light, while eye position was monitored with an infrared camera. High-resolution echo-planar images were acquired with a surface coil in a horizontal-bore 3-Tesla MRI scanner (Siemens Trio), as the monkey was stimulated with the BPN sounds. All data were pre-processed and aligned with the 112RM-SL_T1 rhesus template for identification of cortical fields and were analyzed with a general linear model. Clusters of active voxels were found consistently along the ascending auditory pathways: in the cochlear nucleus, inferior colliculus, medial geniculate nucleus, as well as in the STG. Three clusters of cochleotopic reversals were found in mid-STG (areas MM, A1, ML), rostral STG (RM, R, AL) and, even more rostrally, in areas RT/RTL with the most anterior activation in a region possibly corresponding to RTp. The present results demonstrate that fMRI performed in awake behaving monkeys at 3 Tesla can reliably reproduce the functional organization established with single-neuron recordings. They support and extend previous findings from passive listening-paradigms obtained in vertical 4.7-Tesla magnets. Using these data for functional localization of cortical areas in individual monkeys under various tasks, more complex organizational features of auditory cortex should become identifiable along the cortical hierarchy, which will then again aid the exploration of these areas with microelectrode techniques.
Dual mechanisms for processing pitch in marmosets (*Callithrix jacchus*)
M.S. Osmanski, D.A. Bendor & X. Wang

A fundamental component of both music and speech is pitch, our perception of how high or low a sound is on a musical scale. After more than a century, the exact mechanisms used by the auditory system for pitch extraction are still being debated. Computationally, pitch can be measured using either spectral or temporal information in the acoustic signal. We have examined whether spectral or temporal information is used for pitch extraction by marmosets (a New World monkey species) using both behavioral and neurophysiological techniques. Marmosets trained on an operant conditioning task were able to use temporal information to discriminate pitch for acoustic stimuli with higher order harmonics and lower fundamental frequencies, while only spectral information was used for acoustic stimuli with lower order harmonics and higher fundamental frequencies. Parallel to these behavioral observations, we found that pitch-selective neurons in auditory cortex extract pitch using temporal information for low pitch sounds composed of high order harmonics, whereas spectral information is used for high pitch sounds with low order harmonics. Our data support dual pitch processing mechanisms, whereby both harmonic templates (spectral) and envelope extraction (temporal) are used to compute pitch. [Supported by NIH Grant R01-DC03180 (X.W.)]
Sensitivity to temporal structure in the human auditory system
T. Overath\textsuperscript{1}\textsuperscript{2}, J. H. McDermott\textsuperscript{3}, J. M. Zarate\textsuperscript{1}, D. Poeppel\textsuperscript{1}

1. Department of Psychology, New York University, New York, NY, USA
2. The UCL Ear Institute, University College London, London, UK
3. Center for Neural Science, New York University, New York, NY, USA

Human speech is structured over multiple timescales. Phonemes, syllables and words carry information at scales ranging from a few tens of milliseconds to seconds, respectively (Rosen, 1999). Although cortical sensitivity to amplitude modulation is well-known, it is unclear how the complex temporal structure of speech and other natural sounds is encoded. Inspired by the use of image scrambling to study object recognition (e.g. Grill-Spector & Malach, 1998), we investigated sensitivity to temporal structure by measuring fMRI responses to speech signals scrambled at different timescales. We reasoned that mechanisms for encoding temporal structure might be tuned to the specific structures in familiar natural sounds such as speech, such that the unnatural signals produced by scrambling would evoke a lower response.

Scrambling was accomplished with a “quilting” algorithm similar to those popular in computer graphics, in which pieces of a source image are rearranged and appended to create a new image. We divided a sound signal into fixed-length segments (30-960 ms) and reordered them pseudo-randomly. To minimize segment boundary artifacts, a segment order was selected that approximately matched the segment-to-segment envelope changes in the original signal; segments were then appended using pitch-synchronous overlap-add. The resulting signals shared the spectral structure of the original sound signal, but differed in temporal structure to an extent that depended on the segment duration.

We used a localizer (960 vs. 30 ms) to define fROIs and then probed their responses to other stimuli. The fROIs were located bilaterally in the STS. Their response generally increased with segment length up to at least 480 ms, suggesting sensitivity to syllable-length structure. We found that the effect of scrambling was a) at least somewhat speech-specific (quilts from control sounds with speech-like modulation spectra failed to produce an effect), b) independent of pitch and prosody (quilts from noise-vocoded speech produced the effect), c) independent of lexical access and syllabic familiarity (quilts from familiar and unfamiliar spoken languages produced comparable effects), and d) consistent across and replicable within subjects. The results reveal pronounced sensitivity to temporal structure in human STS that cannot be explained simply in terms of amplitude modulation sensitivity. It remains to be seen whether the effect is fully specific to speech or whether the regions implicated are also sensitive to temporal structure in other natural sounds.

References:
Cochleotopic organization of human Heschl's gyrus as defined by intracranial gamma-band responses

*H. OYA¹, K. V. NOURSKI², H. KAWASAKI², J. F. BRUGGE²,³, R. A. REALE²,⁴, T. D. GRIFFITHS²,⁵, M. STEINSCHNEIDER⁶, M. A. HOWARD, III²;
¹Neurosurg., Univ. of Iowa Hosp. and Clinics, IOWA CITY, IA; ²Neurosurg., Univ. of Iowa Col. of Med., IOWA CITY, IA; ³Physiol., Univ. of Wisconsin, Madison, WI; ⁴Psychology, Univ. of Wisconsin, Madison, WY; ⁵Cognitive neurology, Univ. of Newcastle, Newcastle, United Kingdom; ⁶Neurol., Albert Einstein college of medicine, Bronx, NY

The model that segregates human auditory cortex into 3 main hierarchically organized regions (core, belt and para-belt) has been adopted in various studies in humans. A consistent finding is that core is located within a confined region in the medial portion of Heschl's gyrus. Recent studies suggest that cochleotopic response property is preserved beyond the core cortex and may extend to the lateral superior temporal gyrus. These have also suggested that the tonotopy is not necessarily confined within the HG and not constrained by the gross anatomy of HG; There is still a debate on the tonotopic frequency gradient with respect to the long axis of HG as well as the number of gradients. These issues still remain to be elaborated. To supplement and validate the findings noted above, we pursued this issue by directly recording neural activity in the HG in humans. Subjects are patients with intractable epilepsy who required chronic invasive recordings for their treatment. The gamma-band response from 290 electrode sites in 18 subjects were analyzed. Each brain was registered MNI average-305 template. Information about tone frequency, calculated as mutual information, was found to be encoded in the gamma band response. Regression analysis showed significant posteromedial-anterolateral gradient in gamma-band magnitude. There was also a significant posteromedial-anterolateral preferred frequency gradient (high to low). Regression was significant in the range of 35~55mm in the x-direction and -35~0 mm in the y-direction. Coordinate-rotation analysis suggests the tonotopic axis in the HG is 80-90 degrees rotated in the planum temporale plane.
Speech comprehension requires mapping low level time-frequency representations onto categorical phonetic representations. How the auditory system decomposes complex sounds into elementary acoustic features and then pools these features to form invariant representations of auditory objects is unknown. To study auditory and phonetic representation, we recorded intracranial signals from epileptic patients, using subdural electrode grids placed over left or right superior temporal gyrus (STG). Patients listened passively to phonetically-transcribed English sentences from a variety of male and female speakers. We evaluated the ability of a phonetic-based neural encoding model to account for speech-induced high gamma (70-150 Hz) responses. We compared the predictive power of the phonetic model to that of linear and nonlinear spectro-temporal auditory models. The phonetic model is based on a set of categorical predictors, one for each of 58 distinct consonant and vowel phones from the TIMIT phonetic alphabet. The linear auditory model is based on the spectro-temporal envelope of the stimulus, while the nonlinear model is based on the modulation energy content. The models are fit to neural responses at each electrode site and predictive power is evaluated as the correlation between actual and predicted responses from an out-of-sample validation data set. For specific STG sites, the phonetic model provided equal or better predictions compared to the auditory models, with correlations up to $r = 0.5$. Phonetic tuning in the fitted models exhibited selectivity for consonant-vowel sequences at specific sites. The results suggest that higher order auditory areas in human STG encode both auditory and acoustically invariant phonetic information. The pattern of phonetic tuning may be consistent with a role in lexical recognition.
Spectral and temporal processing of ‘vocoded’ communication signals in the monkey brain

C.I. Petkov\textsuperscript{1} & J. Obleser\textsuperscript{2}

1. Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, U.K.

A prominent hypothesis posits an asymmetry in human auditory processing, such that the left hemisphere specializes in processing the temporal aspects of sound (e.g., speech envelope) and the right hemisphere in the spectral or fine-frequency aspects. It is unclear whether such hemispheric differences are also present in nonhuman animals during spectro-temporal processing of communication sounds. Also, is there a difference in how the auditory cortical fields within a particular hemisphere encode the spectral and temporal features of communication sounds? To address this, we conducted functional magnetic resonance imaging (fMRI) with awake macaque monkeys as they listened to species-specific communications (i.e., exemplars of macaque ‘grunts’ and ‘coos’). The sound stimuli were parametrically degraded (“vocoded”) \cite{1} along two orthogonal dimensions; the temporal/envelope and the spectral/fine-structure detail (5 levels each). We tested for differential sensitivity in brain activation to either of these degradations along the length of auditory cortex on the supratemporal plane. Two awake macaques were imaged as they listened to the experimental stimuli. For analyzing the fMRI data, we used parametric regressors, and, to test for specific activation differences, we evaluated the following fMRI activity contrasts: ‘spectral’ > ‘temporal’ and ‘temporal’ > ‘spectral’. We observed a slightly stronger representation of the spectral dimension in the right hemisphere (stronger temporal representation in the left). This was accompanied by an anterior-posterior processing gradient: Clusters closer to auditory field A1 seemed to more strongly represented spectral features, while clusters anterior to A1, including fields R and RT, more strongly represented temporal features. These macaque fMRI results appear to show an interesting correspondence in a nonhuman primate to a previous study reporting on human hemispheric differences for the encoding of the spectral and temporal features of vocoded speech \cite{2}. Moreover, in this study the greater fMRI activity response to temporal features in anterior auditory cortical regions of a nonhuman primate might be hypothesized to be supported by the temporal processing properties of ‘unsynchronized’ neurons that have been described in the anterior auditory cortex of marmoset monkeys \cite{3}. In summary, the macaque fMRI results help to link human fMRI studies on the spectro-temporal processing of sound to the neuronal level detail that can readily be obtained in nonhuman animals.

\begin{itemize}
\item \cite{1} Shannon RV et al., (1995) Science 270:303-304.
\end{itemize}

Supported by: Max-Planck Society (JO) and Wellcome Trust (CIP).
Sensitivity for sound-movement direction in the macaque auditory cortex.

C. POIRIER, S. BAUMANN, C. I. PETKOV, A. REES, A. THIELE, T. D. GRIFFITHS

Previous human fMRI studies have demonstrated the involvement of the posterior auditory cortex in auditory movement perception (e.g. [1,2]). In a recent fMRI study, we have shown that in macaques the bilateral caudal lateral belt areas (CL) as well as A1 are more activated by moving sounds than by static sounds [3]. The present study sought specific direction sensitivity for movement. Using a primate-dedicated vertical MRI scanner (Bruker 4.7 Tesla), we recorded the BOLD response in the superior temporal gyrus to sounds moving rightward or leftward within each hemifield. Virtual-acoustic space (VAS) stimuli were created for fMRI by recording the pressure waveform within the ear canals of each individual, during the presentation of sinusiodally amplitude modulated broadband noise (rate 80Hz, modulation depth 80 %) from different spatial locations in azimuth (-80 to +80 degrees). To confirm that the monkey could identify the spatial locations of the VAS stimuli, we played them back over headphones and observed that the monkey visually tracked the simulated sound locations. Four moving sound stimuli were then created by concatenating static sounds from -80 to 0 degrees, 0 to +80 degrees, +80 to 0 degrees and 0 to -80 degrees. Control stimuli were static sounds coming from -80, -40, 0, +40 and +80 degrees. Contrasting moving sounds with static sounds confirmed the specific activation of bilateral A1 and CL by moving sounds and extended this motion-sensitive network to the bilateral posterior superior temporal sulcus, in a similar region to macaque visual motion areas MT/MST. Comparing the BOLD response induced by sounds moving in opposite directions revealed direction sensitivity in the same network, namely A1, CL and putative MT/MST. These results support the role of the posterior auditory cortex in specific analysis of sound movement.


Amy Poremba\textsuperscript{1,2}, James Bigelow\textsuperscript{1}, and Ryan Opheim\textsuperscript{1} 1Department of Psychology, Division of Behavioral and Cognitive Neuroscience, University of Iowa, Iowa City, IA 52242, 2Neuroscience Program, University of Iowa, Iowa City, IA 52242.

Neurophysiological recording and lesion studies in non-human primates have demonstrated that the primary auditory cortex (A1) may preferentially respond to pure tones (e.g. Merzenich et al., 1973). This region also responds to complex conspecific vocalizations (Recanzone, 2008) and plays a role during discrimination of these simple and complex sounds (Heffner & Heffner, 1986). The core region of auditory cortex also shows a high degree of neural plasticity during auditory conditioning tasks (Weinberger, 2010; Fritz et al., 2007). Here, we record single-unit neuronal activity from left hemisphere A1 of rhesus macaques on an auditory go/no go delayed matching-to-sample (DMTS) task wherein two acoustic stimuli (500 ms), separated by a fixed memory delay (5000 ms), are either identical sound presentations or two different sound presentations. A small candy reward is delivered after a correct button press on match trials. Each training session consists of 200 trials, where half the trials are match trials and half are nonmatch trials. Eight sound stimuli are chosen from a larger sound set with each sound stimulus repeated 8-12 times on both match and nonmatch trials within an individual recording session. Sound stimuli include pure tones, frequency sweeps, man-made environmental sounds, monkey vocalizations, human vocalizations, natural sounds, synthesized sounds, and other animal sounds. Neuronal firing rates recorded during both task-related activity and passive listening of the same sound stimuli are analyzed.

Neurons were not assessed a priori for preferred frequencies. Single-unit analyses of post-training tuning assessments suggest 76\% of the recorded neurons are tuned to specific frequencies. There were a large percentage of units that responded to sample (44\%) and test stimuli (37\%). Preliminary population activity (n=129 neurons) exhibits significant early increases in the first two seconds of the delay, and late decreases in activity during the last two seconds of the memory delay. These changes in firing rate were predictive of correct behavioral performance on match and nonmatch trials.

Although memory delay related changes are greater in this primary auditory region compared to the dorsal temporal pole and dorsal lateral prefrontal cortex in our laboratory on the same task, the match/nonmatch trial related changes to the test stimuli occur later than in the aforementioned regions. Overall, these results suggest that this region may be involved in the memory encoding of the sample stimulus and/or preparing for the upcoming test stimulus, with a recognition memory component occurring comparatively late during the test stimulus offset.

Support: Startup funds from the Univ. of Iowa and NIDCD grant DC0007156 to A.P.
Modeling the influence of inhibition in shaping temporal coding in the medial geniculate body (MGB)

Cal F. Rabang and Edward L. Bartlett

MGB representations of sound have been shown to transform as information passes from inferior colliculus (IC) to auditory cortex (AC). One example is the segregation of synchronized responses and non-synchronized rate responses to temporally-modulated auditory stimuli arising from synchronized IC inputs. A computational MGB thalamocortical (TC) neuron model, constructed in NEURON, has previously shown that large excitatory inputs exhibiting synaptic depression preserve IC input synchrony up to approximately 50 Hz. Small excitatory inputs exhibiting synaptic facilitation can be synchronized or non-synchronized and display high-pass firing rate responses, filtering out low frequency inputs. These initial models did not include feedforward inhibitory inputs from the IC or feedback inhibitory inputs from the thalamic reticular nucleus (TRN). We test the hypothesis that addition of inhibitory inputs will increase the temporal precision (vector strength) of MGB neuron responses, extend the range of phase-locking by MGB neurons, and sharpen the ranges of effective frequencies coded by rate tuning. We will also test the hypothesis that inhibition that precedes excitation, which would be expected for the IC inhibition, will play a different role than inhibition that follows excitation, which would be expected for the TRN inhibition. We add modeled GABA_A and GABA_B inhibitory synapses to our existing model, using amplitude and decay values fit to data from previous studies (Bartlett and Smith 1999, 2002) and whole-cell recordings (Venkataraman and Bartlett, 2011 SFN abstract). MGB responses are analyzed for synchrony and firing rate relative to the IC input train stimulus or modeled sound.
Our brains must be able to represent sound scenes with a wide range of statistical properties. One important property is the spectrotemporal contrast in the acoustic environment: the variation in sound pressure in each frequency band, relative to the mean pressure. We have shown that neurons in ferret primary auditory cortex (A1) rescale their gain to partially compensate for the spectrotemporal contrast of recent stimulation [Rabinowitz et al., Neuron (2011)]. When contrast is low, neurons increase their gain, becoming more sensitive to small changes in the stimulus.

This work raises two questions. First, does contrast gain control arise in auditory cortex? Second, over what spectral range does a neuron integrate to determine its gain? To answer these, we designed synthetic stimuli with random contrast across different frequency bands, and recorded from neurons in the inferior colliculus (IC) and A1 of anaesthetised ferrets.

We observed contrast gain control in A1 and IC. For units in both regions, the spectral dependence of gain control was similar to units' linear frequency tuning. This means that frequency bands that elicit neural responses also decrease the neuron's gain.

However, the gain control mechanisms showed differences in spectral bandwidth between regions. In IC, gain control had broad frequency tuning (similar to the linear frequency tuning of the units). In A1, gain and linear bandwidths were typically narrow. Furthermore, gain control was stronger in cortex, and had a longer adaptation period (~100ms vs ~10ms). This suggests that separate contrast gain control mechanisms operate in IC and A1.
Effects of silencing parvalbumin-positive interneurons on tone-evoked responses in auditory cortex in vivo

Ashlan P Reid, Tomas Hromadka and Anthony M Zador
Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

In sensory neocortex, inhibition performs many important functions, from providing a brake on recurrent excitation, to aiding in the formation or maintenance of tuning and enforcing precise spike timing. Cortical neurons receive inhibitory input from several diverse populations of local interneurons. Although careful characterization of these different classes of interneurons and their synaptic partners in vitro has yielded important insights, it remains unclear whether different populations of interneurons serve different functions in cortical processing in vivo. This is largely because it is difficult to reliably manipulate the activity of specific neurons in vivo using conventional methods.

Here we describe an optogenetic approach to reversibly silencing the parvalbumin (PV) class of fast-spiking interneurons in the auditory cortex in-vivo. We use the cre/lox system to express the light-activated outward proton pump, Archaerhodopsin-3 (Arch), in PV-ires-cre mice. By targeting expression of AAV-DIO-Arch to PV-interneurons in the auditory cortex, we have achieved both local and fast temporal control of their activity in-vivo. Although we focus here on PV-interneuron mediated inhibition in auditory cortex, the same approach can be readily applied to other neuronal populations, as well as to other sensory areas.

We used in vivo physiological methods to compare responses to single pure tones under normal conditions with responses under conditions in which the PV input was transiently removed via a flash of green light. We recorded local field potentials, spiking responses, and excitatory and inhibitory conductance changes. We found that PV inputs constitutes a major--and under some conditions the sole--source of tone-evoked inhibitory input. Our data suggest a model in which PV neurons receive direct feedforward thalamic input, and this input interacts with local cortical circuitry to shape the sound-evoked response profile.
Human auditory cortex activations to phoneme and nonphoneme vowels during discrimination and memory tasks

T. RINNE, K. HARINEN

Institute of Behavioural Sciences, University of Helsinki, Finland

We used fMRI to investigate how different listening tasks modulate the activations of human auditory cortex (AC) to phoneme and nonphoneme vowels. Subjects (N=16) were concurrently presented with auditory and visual stimulus streams in 15-s blocks alternating with 8-s breaks with no stimuli. The auditory stream consisted of vowel pairs (two 200- ms vowels separated by a 100-ms gap, onset-to-onset interval 900-1100 ms) from three Finnish phoneme (/u/, /a/ and /i/) or three nonphoneme vowel categories. The vowel categories were defined based on a psychoacoustic vowel chart of Finnish vowels and each category contained 9 different vowels. The visual stream consisted of Gabor gratings with varying orientation. In all task blocks, subjects were presented with within-category vowel pairs. In discrimination tasks, they were required to indicate when the first and second part of the vowel pair were the same. In n-back memory tasks, subjects indicated when the vowel pair belonged to the same vowel category as the one presented 1, 2 or 3 trials (depending on the difficulty level) before. Both tasks were performed either on phoneme or nonphoneme vowel pairs in separate blocks (3-back memory task was performed only on phoneme pairs). In the visual task, subjects were required to detect orientation changes. Based on our previous studies, we expected that discrimination tasks would be associated with enhanced activations in the anterior AC, while memory tasks would activate more posterior AC areas. We also expected that these task-dependent AC activations would depend on whether the tasks are performed on phoneme or nonphoneme vowels.

Initial fMRI data analysis showed stronger activations in wide temporal and insular areas during discrimination than during n-back memory tasks, while areas in the inferior parietal lobule (IPL) were more activated during memory tasks. Comparisons between phoneme and nonphoneme blocks during (1-back and 2-back) memory tasks revealed stronger activations during phoneme blocks in an IPL area and stronger activations during nonphoneme blocks in several areas of the anterior and posterior superior temporal gyrus (STG). Similar differences between phoneme and nonphoneme blocks were also observed during discrimination tasks but not during the visual task. These results demonstrate the complex dependence of anterior and posterior AC activations on the characteristics of behavioral tasks and stimuli.
Temporal integration of faces and vocalizations in the ventrolateral prefrontal cortex of non-human primates

Lizabeth M. Romanski\textsuperscript{1} and *Jaewon Hwang\textsuperscript{2}

Neurobiology and Anatomy, University of Rochester School of Medicine, Rochester, NY; \textsuperscript{2} Brain and Cognitive Sciences, University of Rochester, Rochester, NY

Abstract:
During human speech and non-human primate vocalizations, a facial gesture (facial posture shift or mouth opening) generally precedes the generation of a vocalization. By comparing the temporal relationship of the face and vocal components in natural stimuli with the time course of neural responses, we can obtain better understanding of how face and vocal information are integrated over time.

In our study, we trained non-human primate subjects to perform a vocalization discrimination task. They were presented with a species-specific vocalization as a sample and were required to detect a mismatching vocalization in subsequent audio presentations. In some conditions, we presented visual stimuli along with vocalizations such as static face images, face movies, expanding circles and fractal images. We also recorded neural activity of single cells in the ventrolateral prefrontal cortex (VLPFC) while our subjects performed the task, and determined the latencies of neural responses from the onsets of visual and auditory components of the stimuli presented.

In 160 cell-stimulus pairs we estimated the latencies either when the vocalization was presented alone or when the vocalization was presented with face stimuli. In a small number of the cases (10/160, 6%) there was a response only when the vocalization was presented alone whereas 38% (=60/160) had a response to both the vocalization alone and the vocalization presented with face stimuli. In 37 out of the latter 60 cases, the latencies to the vocalization with face stimuli were shorter than that to vocalization alone. It is possible that the integration of the visual information reduced the latency for vocalization processing. Interestingly the modulation of the neural response by the integration of visual stimuli was not always the same as the response to the vocalization alone. For example, in 18 cases out of 60, the response to the vocalizations alone was inhibitory while the response to the vocalization plus faces was excitatory in the same cell.

These various types of responses show the complexity of face-vocalization interactions which occur at the single cell level in VLPFC, and suggest that the temporal aspect of neural activation is an important feature to be considered in the integration of faces and vocalizations which occurs during communication.
Adaptation of neuronal responses to repeated tones in auditory cortex of awake freely moving mice implanted with tetrode arrays

Michael T. Rutledge¹, John O'Keefe¹, and Jennifer F. Linden², ³
⁰ Cell and Developmental Biology, ¹ Ear Institute, ² Neuroscience, Physiology & Pharmacology, University College London, London, United Kingdom

Neurons in the auditory cortex show sensitivity to stimulus statistics over long time scales, from hundreds of milliseconds to many seconds. The mechanisms underlying this long-term adaptation are currently poorly understood. To explore these mechanisms in the genetically tractable mouse model, we are characterizing adaptation of auditory cortical responses to repeated tones in awake freely moving mice, using chronically implanted tetrode arrays.

Mouse brain structure varies significantly with age and across inbred strains (Wahlsten, Hudspeth & Bernhardt, 1975), so we developed an age- and strain-independent method for targeting auditory cortex with chronic implant recording arrays. The target location is determined based on stereotaxic positioning of the temporal ridge and the ventral and rostral squamosa sutures in skull preparations from animals of the same strain and similar age as the animal to be implanted. During surgery, target coordinates are normalized relative to the distance between bregma and lambda, and the implant is placed on the dorsal surface of the skull and angled to access the target location. We demonstrate that this procedure significantly improves auditory cortex targeting accuracy for CBA/Ca mice over that achievable using the standard C57Bl6 mouse brain atlas; it is also faster, easier and cheaper than three-dimensional histological reconstruction of brain tissue.

Tetrode arrays implanted using this method intersect areas with in vivo physiological as well as histological characteristics of auditory cortex. Single-unit and local field potential recordings reveal short-latency tone-evoked responses, clear tuning for tone frequency, and adaptation to repeated tones. Moreover, we find that tone-evoked responses in some neurons show tuned excitation, whilst others show tuned inhibition. In some cases, these two types of response are observed in neighbouring neurons. Excited neurons often show adaptation to repeated tones over hundreds of milliseconds, whilst inhibited neurons generally do not. This observation resembles previous findings in inferotemporal cortex of awake macaque (Sobotka & Ringo, 1994), and suggests that tone-responsive excited and inhibited neurons play different roles in auditory cortical adaptation to long-term stimulus statistics.

Supported by: Wellcome Trust (084364/Z/07/Z and 082507/Z/07/Z), Biotechnology and Biological Sciences Research Council (BB/F0167941/1)
State-dependent interactions in auditory thalamocortical network

Shuzo Sakata
Strathclyde Institute of Pharmacy and Biomedical Sciences
University of Strathclyde
161 Cathedral Street, Glasgow G4 0RE, UK

Internal brain states affect sensory processing and perception, but underlying mechanisms are still elusive. Thalamocortical circuits have been long implicated in both normal and abnormal sensory perception. However, it is still unclear how the thalamus functionally interacts with the sensory cortex in different brain states. Here we investigated state-dependent population activity in the auditory thalamocortical circuit. By using a massively parallel extracellular recording technique, we simultaneously monitored neural population activity in both the primary auditory cortex (A1) and the ventral division of the medial geniculate body (MGv) of rats in vivo. To manipulate brain states, we electrically stimulated the basal forebrain (BF) under anesthesia. First, we assessed the effect of brain states on spontaneous activity: BF activations altered spontaneous firing patterns in both A1 and MGv neurons, with the suppression of cortical superficial layers and less burst firing in most cells. At the population level, BF activations desynchronized neural populations across thalamocortical neurons. During the inactivated state, on the other hand, while the thalamocortical circuit was frequently co-activated, spontaneous firing events (“up states”) propagated from the deep layers of A1 followed by activations of superficial layers and MGv neurons. Second, we characterized state-dependent auditory processing: during the activated (desynchronized) state, temporal tuning was improved in both A1 and MGv. Thalamic activity appeared in synchrony, resulting in rapid propagations of activity from MGv to A1. Moreover, the late component of evoked responses in A1 became more apparent. We hypothesize that cortical activations reduce internal noise and evoked synchrony can recruit neural ensembles globally for sensory processing.
Neural coding of vocalizations in auditory scenes transforms along the auditory pathway
David M. Schneider and Sarah M. N. Woolley

Keywords: cocktail party, songbird, inferior colliculus, auditory cortex

Vocal learners such as humans and songbirds often communicate in distracting acoustic environments. Although psychophysical experiments and everyday experience indicate that the brain suppresses background sounds and/or selectively amplifies target sounds, the neural mechanisms by which these processes occur remain unclear. We investigated how the neural coding of vocalizations changed with varying levels of background sound in the awake zebra finch. We recorded from single neurons in the auditory midbrain (MLd), the primary auditory forebrain (Field L) and a higher auditory region (NCM) while presenting individual zebra finch songs, a chorus of zebra finch songs, and auditory scenes consisting of a single song and the chorus. Auditory scenes were presented with varying signal-to-noise ratios (SNRs) by systematically changing the song intensity between +/- 15 dB while keeping the chorus intensity constant.

Neurons in all three brain areas responded to vocalizations, but neurons in the higher forebrain fired far fewer spikes and were usually selective for a subset of vocalizations. Midbrain and primary forebrain neurons fired strongly in response to chorus and auditory scenes, with scene firing rates being largely SNR-invariant. Alternatively, many higher forebrain neurons completely stopped responding during the playback of chorus, and their firing rates to auditory scenes was strongly SNR-dependent, such that these neurons rarely responded to low-SNR scenes.

In all three brain areas, the spike train patterns evoked by auditory scenes transitioned between being chorus-like and song-like as the SNR increased. At successive stages of the auditory pathway, this transition differed in three important ways. First, at high SNRs, the spike trains of auditory forebrain neurons were substantially more song-like than the spike trains of auditory midbrain neurons, indicating a hierarchical improvement in the filtering of background sound. Second, the SNR at which neurons transitioned between chorus-like and song-like spike trains decreased at successive processing stages, indicating a hierarchical shift in the neural extraction of vocalizations from auditory scenes. Third, in the auditory forebrain, and particularly the higher forebrain, neurons produced more step-like transitions between chorus-like and song-like spike trains, indicating a hierarchical emergence of categorical responses. Together, these results show a marked change in the neural coding of auditory scenes at successive stages of the auditory pathway, and particularly between primary and higher forebrain areas.
Half of the members of the multigenerational KE family have an inherited speech-language disorder, characterised mainly as a verbal and orofacial dyspraxia, caused by a mutation of the FOXP2 gene. The core phenotype of the affected KE members (aKE) is a deficit in repeating words, especially non-words, and in moving the orofacial musculature. It is not clear whether the speech deficit results from combined Working Memory (WM) and motor output problems, or from motor output difficulties alone.

WM was investigated using a test battery based on the Baddeley and Hitch WM model. The model posits that the central executive (CE), important for planning and manipulating information, works in conjunction with two modality-specific components: The Phonological Loop (PL), specialized for processing speech-based information, and the Visuospatial Sketchpad (VSSP) dedicated to processing visual and spatial information. We compared WM performance related to PL, VSSP and CE function in aKE (N=3) and healthy controls (N=5, 2 unaffected KE members).

An ANOVA examining the effects of Component (PL, VSSP, and CE) and Group (aKE and controls) revealed only a significant Component*Group interaction. Compared to controls, the aKE members showed a trend toward lower performance on PL but not on VSSP or CE. Furthermore, the aKE members were impaired relative to controls in the recognition-based PL subtest (word list matching), where repetition (i.e. motor output) of speech-based material was not required.

These preliminary results indicate that the aKE, who have both structural and functional abnormalities in Broca’s area, a structure known to be involved in phonological WM, may be specifically impaired in phonological WM, but not in CE or VSSP. This suggests that the word and non-word repetition difficulties of the aKE members may be influencing their motor-related representations required for internal rehearsal of speech-based material in phonological WM.

References


Subcortical connections of the supratemporal plane and rostral superior temporal gyrus in macaque monkeys.

Laboratory of Neuropsychology, NIMH, NIH, Bethesda, MD 20892 The flow of information in the auditory cortex follows a hierarchical processing scheme, with projections in the medial-lateral dimension from core, to belt, to parabelt (Hackett, Hear. Res. 2010). We recently showed that rostral auditory cortex receives input via stepwise serial projections in the caudal to rostral dimension: through the primary, rostral, and rostrotemporal core fields (AI, R, and RT) in the supratemporal plane, continuing to the rostrotemporal polar field RTP (Scott et al., Soc Neur Abs, 2010). In this study we describe the subcortical connections of AI, R, RT, and RTP, as well as the rostral parabelt (RPB) and the rostral superior temporal gyrus (STGr) in rhesus monkeys. We placed injections of retrograde tracers Fast Blue, Diamidino Yellow, or Cholera toxin b subunit in AI, R, RPB and STGr. Injections of bi-directional tracer Fluororuby or the anterograde tracer biotinylated dextran amine were placed in RT and RTP. Within the core, the balance of thalamocortical input shifts between the caudal and rostral core fields: AI receives a preponderance of thalamic input from the ventral division of the medial geniculate (MGv), R receives a mix of inputs from both the ventral and dorsal subdivisions (MGv and MGd, respectively), and inputs to RT skew more heavily towards MGd. Fluororuby injections into RT and RTP resulted in dense anterogradely labeled terminals in MGd and MGv but few retrogradely labeled neurons in both subdivisions, indicating strong corticothalamic projections. Injections beyond the core fields in RTP, RPB, and STGr revealed thalamic inputs from outside the medial geniculate, including the suprageniculate nucleus and medial pulvinar. Anterograde label was also evident in the striatum: RT and RTP project to the tail of the caudate nucleus and adjacent ventral putamen, with RTP also projecting to the ventromedial region of the head of the caudate nucleus. Finally, RT and RTP are strongly interconnected with the amygdala, as evidenced by extensive anterograde label in the lateral amygdala, and both antero- and retrograde label in the accessory basal nucleus. In summary, auditory fields rostral to the core receive little input from the auditory thalamus, suggesting that their physiological responses to sound are mediated by the previously described corticocortical pathways. The subcortical connectivity described in this study establishes circuits by which auditory processing may interact with multisensory information from the thalamus, as well as systems involved in learning, reward, and emotion (i.e., the striatum and amygdala).

Supported by NIMH Intramural Research Program
Auditory working memory in rhesus monkeys Scott, B.H.1, Yin, P.2, and Mishkin, M.1 1Laboratory of Neuropsychology, NIMH, NIH, Bethesda, MD 20892 2Institute for Systems Research, Univ. of Maryland, College Park, MD

The aim of the present study was to demonstrate auditory working memory in a non-human primate, and to determine the capacity, specificity, and duration of the maintained representation. We trained two monkeys on a serial delayed-match-to-sample task analogous to that used in prior studies of visual memory. The animal initiated a trial by holding a touch bar, and was presented with a sample stimulus which was followed by 1-3 test stimuli at an inter-stimulus interval of ~1 s. Bar release following a sound that was identical to the sample (a match) earned a liquid reward, whereas bar release following a nonmatch sound was recorded as a false alarm (FA), which aborted the trial and initiated a punitive time-out. Match and nonmatch stimuli were drawn randomly from a set of 21 sounds, each ~300 ms in duration. Stimuli included 3 exemplars from 7 categories, including synthesized sounds (rippled noise, band-pass noise, pure tones, frequency-modulated sweeps), and recorded natural sounds (rhesus vocalizations, other animal vocalizations, and environmental sounds). Performance of both monkeys was accurate in the absence of distracters, but degraded severely after the presentation of intervening nonmatch stimuli. Manipulation of the inter-stimulus interval confirmed that the performance degradation was attributable to the appearance of the intervening stimuli, not the decay of memory over time. Analysis of performance by stimulus type showed a weak and counter-intuitive effect of sound category that favored simple stimuli (e.g., tones and noise) and not species-specific vocalizations. The probability of FA occurrence following a nonmatch sound was positively correlated with the spectral similarity of the nonmatch sound and the preceding sample, but this correlation became weak following an intervening nonmatch. Our analysis suggests a strong effect of retroactive interference, such that the nonmatch stimulus disrupted the maintained trace of the sample, effectively lowering the similarity criterion at which monkeys indicated a match. The present results suggest that the auditory working memory trace in nonhuman primates is fragile, and limited to one or two items. The difference in performance relative to visual working memory may reflect a less robust mechanism, and/or a difference in strategy whereby stimulus category is not exploited.

Supported by NIMH Intramural Research Program
Comparison of invasive depth electrode and magnetoencephalographic virtual electrode recordings of induced gamma responses to pitch

William Sedley, Sundeep Teki, Sukhbinder Kumar, Tobias Overath, Gareth Barnes and Tim Griffiths

We have previously recorded oscillatory responses to regular interval noise (RIN) pitch stimuli from depth electrodes along the axis of Heschl’s gyrus (HG), suggested to be neural correlates of the pitch percept (Griffiths, Kumar, Sedley et al, Direct recordings of pitch responses from human auditory cortex, Curr. Biol. 22 (2010), pp. 1128-1132). In the present study we compared these responses with those from ‘virtual electrodes’ based on magnetoencephalography (MEG) recordings in normal subjects. The MEG work allows a further test of the hypothesis that gamma band responses to pitch are correlates of pitch arising from HG, in addition to allowing virtual electrode placement in locations and subjects where invasive electrode placement is not possible. We carried out MEG recordings using a 275 channel whole-head MEG setup with third order gradiometers (CTF systems) from 13 healthy volunteers. Dynamic imaging of coherent sources (DICS) beamformer analysis was carried out to produce brain-wide power maps and virtual electrode spectrotemporal profiles from auditory cortex. In keeping with the depth electrode recordings, oscillatory activity during pitch perception was maximal in the medial and middle parts of HG, corresponding to primary and adjacent auditory cortex. Time-frequency decomposition of actual and virtual depth electrode data yielded concordant results in the form of induced gamma band oscillations from around 60 Hz upwards occurring from around 70ms after transition from control noise to RIN and persisting for the duration of the RIN. The striking concordance of the anatomical locations, spectrotemporal profiles and power magnitudes of the two data types supports the validity of this approach to accurately simulate invasive electrode recordings using MEG.
Change of frequency tuning underlies stimulus-specific adaptation in the inferior colliculus of rat

Li Shen, Lingyun Zhao, *Bo Hong
Department of Biomedical Engineering, Tsinghua University, Beijing, China
Email: bmeshenli@gmail.com, hongbo@tsinghua.edu.cn

Abstract
Stimulus-specific adaptation (SSA), in which rare stimuli elicit higher responses than common stimuli, was observed at single neuron level in auditory system (Ulanovsky, 2003; Malmierca, 2009). However, it remains unknown how receptive field changes after adaptation and how much these changes relate to SSA. We explored this question by comparing frequency tuning curves of neurons in rat inferior colliculus in an adapted condition and a normal condition. Normal frequency tuning curves were measured at 20 frequency points centered at neuron’s best frequency. Adapted tuning curve was measured by the same set of frequency probes as in the normal condition, except they were randomly interspersed in a repeating tone sequence with a fixed frequency which accounted for 90% of the sequence. We found that after adaptation, the responses to frequencies very close to the repeated one were significantly decreased while to frequencies at farther vicinities showed increased responses. The difference between normal and adapted tuning curves revealed the particular part of input to the neuron that was suppressed by adaptation. This part of input showed a lateral inhibition. We also stimulated neurons using oddball paradigm with the same common stimuli and found the responses to rare ones can be largely predicted by the adapted tuning curve. This result suggests that neurons’ receptive field undergoes elaborate dynamics during adaptation to the ongoing sound stream (Malone, 2001; Rahne, 2009), which may enhance information transmission in adapted condition. Furthermore, it provided evidence of the neural mechanism of SSA and suggested more properties of adaptation which were not shown by current SSA experiments. (This project is supported by National Program on Key Basic Research Projects of China 2011CB933204)

References
Brainstem correlates of pattern learning

Erika Skoe, Emily Spitzer, Nina Kraus

The ability to extract patterns from continuous sound sequences is fundamental to music as well as native and second language learning. In light of recent research showing that auditory brainstem activity is sensitive to simple, highly-repetitive patterns and stimulus context, we seek to elucidate if (and how) subcortical structures are involved in this generalized form of pattern detection, the time course over which learning-related subcortical plasticity can be observed, and the degree to which behavioral and neural measures of pattern detection are linked to performance on a commonplace, complex auditory task like listening to speech in background noise. We have chosen to examine speech in noise perception because real-world listening conditions are generally noisy, and because pulling out a signal from noise depends on the ability to learn the patterns associated with the target and the noise. In a group of normal hearing young adults, we evaluated the auditory brainstem response to eight complex tones presented in a patterned and an unpatterned sequence. In the patterned condition, the eight tones were grouped into four doublets, such that the tones forming a doublet were always presented in immediate succession. In the unpatterned condition, the eight tones were presented in pseudo-random order. In both conditions, the time interval between tones was identical and each tone had a 12.5% probability of occurrence; what differed between conditions was the transitional probability between successive tones. We measured the auditory brainstem response to the fundamental frequency of each tone and calculated the extent to which the response changed between the patterned and unpatterned conditions. Preliminary results show a relationship between temporal discrimination ability and neural pattern enhancements, such that subjects with greater neural enhancements between conditions were better at discriminating two sounds occurring in rapid succession, as assessed by a backward masking task. Our results suggest that poor temporal resolution might compromise the brain’s ability to segment the tone sequence into temporally distinct patterns. Our findings are noteworthy given that backward masking thresholds, which were found to correlate with the ability to hear speech in noise, are (1) elevated in children with
language impairments, a group demonstrating impaired speech perception in noise and impaired pattern learning; and (2) lower in musicians, a group with exceptional perceptual abilities in noise.
In the present study we measured the spectral processing underlying auditory spatial tuning in nontonotopic regions of the inferior colliculus in unanesthetized marmoset monkeys. Using stereotaxic landmarks, we attempted to target the nucleus of the brachium of the IC (BIN), which provides a major auditory projection to the superior colliculus and has been implicated in the formation of the auditory space map present there. Sound localization cues including interaural time differences (ITDs), interaural level differences (ILDs), and spectral shapes (SSs) were measured in each monkey. The results were used to filter broadband noise to create virtual space stimuli. In addition we synthesized binaural random spectral shape (RSS) stimuli, which had random and independent levels between each ¼ octave frequency bin and each ear so as to create ILDs in each bin with a standard deviation of 12 dB across the stimulus set. Both stimuli as well as pure tones were presented in close field while recording single unit spike responses in the IC. The results can be summarized as follows: 1) Neurons in nontonopic IC were usually broadly tuned in frequency with onset, sustained, or no responses to tones. 2) Spatial receptive fields were usually tuned to the contralateral hemifield. Some neurons were broadly tuned while others showed relatively sharp tuning in both azimuth and elevation. 3) Responses to binaural RSS stimuli were used to constrain models that transform the binaural stimulus spectrum into discharge rate using a combination of first- and second-order weighting of the spectrum. The first-order (linear) weights were often broad, excitatory in the contralateral ear, inhibitory in the ipsilateral ear, and asymmetric across ears. Virtual space stimuli were then used as inputs to the models and compared to the neural responses. In some cases the first order RSS model accurately predicted the measured receptive fields while in other cases second-order weighting functions led to significant improvements. The results suggest that both linear and nonlinear spectral processing underlies spatial tuning in nontonotopic regions of the IC.
Neurons in the auditory cortex (AC) are simultaneously selective (“tuned”) along multiple stimulus dimensions, the primary dimensions being tone frequency and intensity at the left and right ears. At each frequency, tuning to intensity at the two ears produces receptive fields that can be characterized along two orthogonal dimensions: average binaural level (ABL) and interaural level difference (ILD, a salient cue for sound localization) [Kitzes LM, *Hear Res* 248:68-76. 2008.]. In this study, functional magnetic resonance imaging (fMRI) was used to similarly characterize the sensitivity of evoked responses in human AC to the ABL (ranging 55-85 dB SPL) and ILD (ranging +/-30 dB) of narrowband-filtered impulse trains centered at 4000 Hz. Two imaging experiments, one using a sparse (12s TR) block design and one using a continuous (2s TR) event-related design, were conducted at 3T (Philips Acheiva). Both experiments parametrically manipulated ILD across stimulus presentations.

Consistent with population responses observed in animal studies, response-ILD functions revealed tuning to contralateral ILD when ABL was held constant. The dynamic range of response was greater across values of ILD than for corresponding intensity variation at the contralateral (preferred) ear, as a result of binaural suppression that was strongest for ILD favoring the ipsilateral ear by 10-20 dB. As a result, response-ILD functions exhibited an asymmetric “U-shape,” where responses to large (30 dB) ipsilateral ILD were similar to monaural responses and in excess of expectations based on contralateral tuning.

Two possible accounts of these data are considered. First, psychophysical models of the binaural summation of loudness [Zwicker E and Zwicker UT, *J Acoust Soc Am* 89:756-65. 1991; Moore BCJ and Glasberg BR, *J Acoust Soc Am* 121:1604-12. 2007] are used to test the hypothesis that AC responses follow the perceived loudness (a psychological dimension which may not be constant for fixed physical ABL) of a sound independent of its location (ILD). Second, neural-population models of AC [Stecker GC et al., *PLoS Biol* 3:e78. 2005] are used to test the hypothesis that responses of ipsilaterally tuned subpopulations, masked at moderate ILD values, contribute strongly to the overall response when extreme ILDs (e.g., monaural sound) are presented. As a consequence, any contralateral preference observed in monaural responses likely underestimates the actual sensitivity to binaural spatial cues such as ILD. Work supported by NSF IOB-0630338, NIH R03-DC009482-02S1.
Tonotopic organization of the human lateral superior temporal gyrus: Implications for complex sound processing

Mitchell Steinschneider1, Kirill V. Nourski2, Hiroyuki Oya2, Hiroto Kawasaki2, Matthew A. Howard III2

1 Department of Neurology, Albert Einstein College of Medicine, New York 10461, NY, USA
2 Department of Neurosurgery, The University of Iowa, Iowa City 52242, IA, USA

Intracranial recordings reveal that posterior lateral portions of the superior temporal gyrus (PLST) are responsive to a broad array of sounds. Ongoing work from our group (see companion abstract by Nourski et al.) provides evidence that PLST is tonotopically organized and may harbor multiple distinct fields. Several lines of investigation stem from these observations. First, is this organization intensity-invariant, or do response patterns markedly diverge when tones of different intensities are presented? Second, can patterns based on pure-tone responses predict differential sensitivity to more complex stimuli, and are there regions of PLST unresponsive to tones but activated by more complex stimuli, thus suggesting higher order organizational schemes?

Experiments were carried out in patients being evaluated for surgical remediation of medically intractable epilepsy. The patients were implanted with high density subdural grid electrodes centered over perisylvian cortex. All procedures were IRB and NIH-approved, and all patients gave informed consent prior to their study participation. Multiple stimulus types were presented, including pure-tone stimuli (250-8000 Hz) of several different intensities, bandpass noise (BPN) bursts with center frequencies matching those of the pure tones, trains of acoustic clicks (rates 25-200 Hz), and consonant-vowel (CV) syllables. The principal measure was high gamma event-related band power.

Results suggest that while there is an expansion of PLST activation with increase in tone intensity, tonotopic features persist. BPN similarly respected basic tonotopic patterns defined by pure tone responses, despite BPN stimuli having one-octave bandwidths. Responses to click trains showed a tonotopically-arranged profile, wherein lower rates were maximal at sites most responsive to high frequency tones, while higher train rates best activated sites most responsive to lower frequency tones. CV syllables elicited the most complex relationships with pure tone responses, but evidence supports some degree of tonotopically-driven patterns when comparing responses varying along consonant place of articulation.

We conclude that within PLST, spectral specificity of tone-evoked responses is an important factor in determining responses elicited by more complex sounds. Ongoing analyses and data collection from additional subjects will refine the associations observed between PLST activation patterns elicited by simple and complex sounds.

Supported by NIH RO1-DC004290, UL1RR024979, and the Hoover Fund.
Current-Source Density and Multiunit Analysis across Layers of Primary Auditory Cortex following Systemic Salicylate Administration in the Rat

Daniel Stolzberg, Richard Salvi, and Brian Allman

Center for Hearing & Deafness, Univ. at Buffalo, the State University of New York

Systemic salicylate administration at high doses is known to induce temporary tinnitus in rats and humans. While previous investigations have sought to characterize salicylate-induced alterations in primary auditory cortex (A1), little information currently exists regarding cortical layer-specific alterations in neural activity. In order to investigate layer-specific changes associated with pharmacologically-induced (250 mg/kg, IP Na\(^+\)-salicylate) tinnitus, we used 32-channel linear array electrodes to sample sound-driven local field potential and multiunit activity simultaneously. Following salicylate, current-source density (CSD) analysis based on the noise-burst evoked translaminar local field potentials revealed significantly enhanced activation (sinks) of granular and supragranular layers. Furthermore, multiunit responses to noise bursts, which were normally sparse in supragranular layers, became apparent following salicylate treatment. In addition, frequency-receptive fields were generated from multiunit activity and the average rectified current (AVREC) of the CSD profile. Multiunit receptive fields shifted and/or expanded their representation to the previously identified pitch of salicylate induced tinnitus in rats (10 – 20 kHz), indicating a disproportionately large representation of the tinnitus pitch along A1 tonotopy. Interestingly, AVREC receptive fields exhibited enhanced responses concentrated to tones near the previously estimated tinnitus pitch and near hearing threshold. Evidence from this study supports the hypothesis that salicylate directly disinhibits A1 resulting in increased sensitivity to sound near the tinnitus pitch and intensity. Furthermore, significantly altered CSD and multiunit activity in supragranular layers implicates intracortical mechanisms may be co-opted for abnormally large spectral integration tuned to the tinnitus pitch. Understanding alterations in A1 neural activity in context of its microcircuitry will hopefully lead to pharmacological interventions for chronic tinnitus in humans.

This research was supported in part by National Institutes of Health grants R01DC0090910, R01DC009219, and F31DC010931-01. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institute on Deafness and Other Communication Disorders or the National Institutes of Health.
Using optimal experimental design for capturing parameters of neural networks in the inferior colliculus of the common marmosets.

AUTHOR BLOCK: *W. TAM1, E. DEKEL1, C. DIMATTINA2, E. D. YOUNG1, K. ZHANG1; 1Johns Hopkins Univ., BALTIMORE, MD; 2Case Western Reserve Univ., Cleveland, OH

Abstract:
Central Nucleus of Inferior Colliculus (CNIC) neurons have traditionally been classified from their responses to basic auditory stimuli, such as response maps (RM) or spectro-temporal receptive fields (STRF). However, due to the nonlinearity of most CNIC neurons, it is beneficial to classify neurons based on neural networks, which are universal function approximators. Neural networks have the abilities to account for the existing classifications and to model nonlinear sensory neurons. However, a large number of stimuli are needed to capture the parameters of these neural architectures. In this study, we employed an optimal experiment paradigm to capture the parameter values of CNIC networks. This computational method reduces the number of stimuli needed for the accurate recovery of nonlinear neural responses by adaptively generating stimuli that optimize the parameters estimation of the hypothesized models (DiMattina and Zhang 2011). We present a family of neural networks and use the optimal experimental design to search for the minimal neural processing necessary to account for properties of the various CNIC units. The family of networks contains an auditory nerve (AN) model at its front-end to mimic the auditory periphery. The outputs from the AN model enters a layer of one, two or three inhibitory neurons that inhibits the eventual principle CNIC unit. In this family of CNIC models, the computationally complicated AN model in the front-end of the network makes the direct on-line optimization and generation of stimuli arduous and time consuming, and therefore the algorithm was modified to choose the optimal stimulus based on estimated AN fiber input to the CNIC networks. A set of such matrices was created off-line, each representing stimuli around a different frequency which accounts for different possible best frequencies (BF) of CNIC neurons. In order to validate the efficiency, accuracy and robustness of the optimal design algorithm, different sets of parameters were chosen from parameters fitted to prior physiological data and the experimental process was simulated with CNIC networks using these values. The simulation results show that the optimal design algorithm can use less than 300 stimuli for recovering the parameters of the neural architecture. The final recovered neural network can potentially predict responses of various CNIC units.
Auditory figure-ground segregation using a complex stochastic stimulus

S. Teki¹, M. Chair², D. Williams³, A. Siddiq², N. Barascud², S. Kumar¹,³, Shamma, S.A.⁴, T. D. Griffiths¹,³

¹ Wellcome Trust Centre for Neuroimaging, University College London, UK
² UCL Ear Institute, University College London, UK
³ Newcastle Auditory Group, Medical School, Newcastle University, UK
⁴ Institute for Systems Research, University of Maryland, College Park

In contrast to the disordered acoustic environment, most studies of auditory segregation have used relatively simple signals. We developed a new stimulus – “stochastic figure-ground” (SFG; Teki et al., 2011) that incorporates stochastic variation in frequency-time space that is not a feature of the predictable sequences used previously. Stimuli consist of a sequence of 50 ms chords containing a random number of pure-tone components. Occasionally, a subset of tonal components repeat in frequency over several consecutive chords, resulting in a spontaneous percept of a “figure” popping out of a background of varying chords. Stimuli were 2000 ms long, with “figures” of duration varying from 100 - 350 ms inserted around 1000 ms post stimulus onset. Figure and background signals overlap in spectrotemporal space, but vary in the statistics of fluctuation and figures can therefore, only be extracted by integrating the patterns over frequency and time. Our behavioral results demonstrate that human listeners are remarkably sensitive to the emergence of such figures (Experiment 1).

To characterize the brain mechanisms that underlie this sensitivity, we sought to investigate the degree to which performance is affected by systematic stimulus manipulations. In Experiment 2, we demonstrate that figure-detection is unaffected when white noise (50 ms) is inserted between successive chords. In Experiment 3, figures were “ramped” (successive figure components were not repeating but increasing in frequency in steps of 2*I or 5*I, where I = 1/24 octave is the resolution of our frequency-bank). Results show decreased sensitivity, although, remarkably, listeners could still perform the task. Experiment 4 tested figure-detection by removing the “background-only” chords, which preceded and followed the figure. Results demonstrated no effect on performance.

Overall, the notable sensitivity exhibited by listeners cannot be explained by existing adaptation-based models of segregation. However the data are consistent with the temporal coherence model of Shamma et al., 2010.

References:


Altered perception after noise exposure during development

Univ. of Virginia, CHARLOTTESVLE, VA

Sensory experience during development is known to influence the organization of the sensory cortices. Changes include modifications at the level of single neurons as well as at population representations. How would these neuronal changes at the sensory level affect information processing downstream? In this report, we examined the behaviors of adult rats reared under noise exposure during development. Infant rats were exposed to pulsed-noise environments after ear-opening until weaning. We measured their behavioral sensitivity to tone and gap detection using acoustic startle and prepulse inhibition to acoustic startle paradigms. We found that noise-reared animals had lower intensity thresholds for a startle response. They also had reduced prepulse inhibition compared to age-matched animals. In addition to an increased behavioral sensitivity of stimulus intensity, noise-reared animals had a different behavioral response to stimulus gaps. Specifically, we observed a reduced inhibition when the duration of the gap matched the duration of the inter-pulse intervals of the noise heard during development. These results showed that noise exposure during development changed animals’ loudness perception of noise. Familiar temporal structure reduced prepulse inhibition to acoustic startle, suggesting an enhanced loudness perception of the sound the animals experienced during development.
The ability to recognize individuals is imperative for social and vocal interactions in all gregarious species. Amongst sophisticated animals, such as nonhuman primates, this process is particularly complex as individuals can be categorized along multiple social dimensions, such as individual, sex, social group, etc. Previous work in common marmosets (Callithrix jacchus) indicated that these social categories are encoded within the spectro-temporal acoustic structure of their phee calls and affect the temporal dynamics of a natural vocal behavior known as antiphonal calling. This natural, species-typical vocal behavior involves the reciprocal exchange of phee calls between visually occluded conspecifics. The aim of this study was two-fold. First, we sought to develop a methodological technique that would allow us to investigate social category perception during antiphonal calling in common marmosets. And second, we tested whether marmosets attend to social categories while engaged in this natural behavior. To this end, we developed a playback method that took advantage of the fact that antiphonal calling typically occurs in bouts of reciprocally exchanged vocalizations. The method we developed utilized custom software to first engage subjects in antiphonal calling and after a predetermined number of exchanges in the bout, we would broadcast a probe or control stimulus. We then measured two responses from subjects - whether they produced an antiphonal call in response to the stimulus (Y/N) and the latency to that response – and compared between probe and control stimuli recorded over the session. Employing this methodology, we tested whether marmosets were able to recognize the individual identity of conspecifics during antiphonal calling. Analyses showed that marmosets produce fewer antiphonal call responses in probe stimuli. Moreover, when subjects do produce an antiphonal call response in this context, they do so at longer latencies. This suggests that marmosets not only recognize individual identity during antiphonal calling, but that this aspect of social categorization can affect the dynamics of these natural antiphonal calling interactions. More broadly, these results suggest that this playback method can be applied to more diverse experiments examining social categorization at both the behavioral and neural levels.

This work supported by grants from the NIH R00 DC009007
Differential representation of speech sound categories between cell classes in the primate superior temporal gyrus

Joji Tsunada¹, Jung H. Lee¹, Yale E. Cohen¹

¹Department of Otorhinolaryngology: Head and Neck Surgery, University of Pennsylvania School of Medicine

The ability to categorize auditory stimuli into meaningful categories is a fundamental cognitive function. To understand the neural computations underlying auditory categorization, it’s essential to examine neural processing not only across cortical areas but within local microcircuits. We have demonstrated how neural representation of auditory category is shaped across cortical areas: neural categorization of auditory stimuli emerges in the superior temporal gyrus (STG) and the latter stage of auditory processing (the ventrolateral prefrontal cortex) represents decision-related signals associated with auditory-category judgments. In contrast, the neural processing of auditory categories within local microcircuits remains unclear. Local microcircuits are constituted by multiple cell types; the main elements of cortical circuits are excitatory pyramidal neurons and inhibitory interneurons. Since recent studies have demonstrated that pyramidal neurons and interneurons play differential or complementary roles in processing visual information, we hypothesized that pyramidal neurons and interneurons in the auditory cortex may differentially process information regarding a sound’s category. To test this hypothesis, we recorded single-unit activity from rhesus monkeys’ STG during a speech-sound categorization task. Next, based on a neuron’s spike-waveform shape, neurons were
classified into one of two categories. In one class of neurons, the spike waveform had a short trough-to-peak time (i.e., narrow-spiking putative interneurons; NS neurons). A second class of neurons had a long trough-to-peak time (i.e., broad-spiking putative pyramidal neurons; BS neurons). We found that NS neurons were more selective for a stimulus’ categorical membership than BS neurons. Moreover, an analysis of these neurons’ temporal profiles indicated that NS neurons showed categorical selectivity earlier than BS neurons. Together, these data indicate that NS neurons and BS neurons in the STG are differentially involved in the auditory categorization and that category specificity is computed through a microcircuit involving NS and BS neurons.
The MGB receives afferent projections from the inferior colliculus (IC), consisting of glutamatergic and GABAergic inputs. The MGB is the only sensory thalamus to receive a lemniscal feed-forward inhibitory input, and the function of this IC-MGB GABAergic synapse is unknown. Given the importance of inhibition in temporal processing, we have studied the development of the IC-mediated inhibitory response in MGB neurons, with a focus on temporal synaptic properties. Additionally, we also studied temporal intrinsic properties of developing MGB neurons by injecting depolarizing current pulse trains at a range of frequencies, in order to observe the intrinsic limitations in temporal processing at different ages. MGB intrinsic and synaptic properties were compared in rats shortly before and after hearing onset (which occurs around P12) at ages P7-9 and P15-17 respectively, and an older control group (P22-32). Recordings from MGB neurons were made using the whole cell patch clamp technique on horizontal rat brain slices that preserved afferent IC-MGB axons. Inhibitory post synaptic potentials (IPSPs) were isolated after pharmacologically blocking glutamate receptors.

IPSPs were present at all age groups tested, and both GABA_A and GABA_B responses could be isolated across ages. Preliminary results suggest that while IPSP amplitudes were comparable across ages, IPSP latency and rise time were shorter in the older animals. In response to repetitive stimulation, IPSPs depressed more in younger animals. In response to current pulse trains at 50 Hz and higher frequencies, spike probabilities increased with age, suggesting greater reliability of thalamocortical relay spikes in older animals at rapid rates. Overall, these results suggest that MGB intrinsic and synaptic responses to repetitive stimuli develop with age, hence providing differential auditory inputs to developing cortical neurons.
Changing microcircuits in the subplate of the neonatal cortex

S. VISWANATHAN\textsuperscript{1}, J. P. Y. KAO\textsuperscript{2}, P. O. KANOLD\textsuperscript{1}

1 Biol., Univ. of Maryland, College Park, MD;
2 Sch. of Med., Univ. of Maryland, Baltimore, MD

Subplate neurons (SPNs) are a population of neurons in the mammalian cerebral cortex that exist predominantly in the prenatal and early neonatal period. Loss of SPNs prevents the functional maturation of the cerebral cortex. SPNs receive subcortical input from the thalamus and relay this information to the developing cortical plate and thereby can influence cortical activity in a feed-forward manner. Little is known about potential feedback projections from the cortical plate to SPNs. Thus, we investigated the spatial distribution of intracortical synaptic inputs to SPNs in vitro in mouse auditory cortex by photostimulation. We find that SPNs fell into two broad classes based on their distinct spatial patterns of synaptic inputs. The first class of SPNs receives inputs from only deep cortical layers while the second class of SPNs receives inputs from deep as well as superficial layers including layer 4. We find that superficial cortical inputs to SPNs emerge after P6 and that SPNs that receive superficial cortical input are located more superficially than those that do not. Our data suggest that distinct circuits are present in the subplate and that while SPNs participate in an early feed-forward circuit they are also involved in a feedback circuit at older ages. Together our results show that SPNs are tightly integrated into the developing thalamocortical and intracortical circuit.
We can recognize the melody of a familiar song when it is played on different musical instruments. Similarly, an animal must be able to recognize a warning call whether the caller has a high-pitched female or a lower-pitched male voice, and whether they are sitting in a tree to the left or right. This type of perceptual invariance to “nuisance” parameters comes easily to listeners, but it is unknown whether or how such robust representations of sounds are formed at the level of sensory cortex. In this study, we investigate whether neurons in both core and belt areas of ferret auditory cortex can robustly represent the pitch, formant frequencies or azimuthal location of artificial vowel sounds while the other two attributes vary. We found that the spike rates of the majority of cortical neurons that are driven by artificial vowels carry robust representations of these features, but the most informative temporal response windows differ from neuron to neuron and across five auditory cortical fields. Furthermore, individual neurons can represent multiple features of sounds unambiguously by independently modulating their spike rates within distinct time windows. Such multiplexing may be critical to identifying sounds that vary along more than one perceptual dimension. Finally, we observed that formant information is encoded in cortex earlier than pitch information, and we show that this time course matches ferrets’ behavioral reaction time differences on a go/no-go change detection task.
Auditory thalamic neurons show nonlinear sensitivity to stimulus context.
R.S. Williamson (1,2), L.A. Anderson (3), G.B. Christianson (3), M. Sahani (1), J.F. Linden (3,4)
1. Gatsby Computational Neuroscience Unit, University College London
2. Centre for Mathematics and Physics in the Life Sciences and Experimental Biology, University College London
3. Ear Institute, University College London
4. Department of Neuroscience, Physiology, and Pharmacology, University College London

Abstract
How does the auditory system achieve its remarkable capacity to operate in acoustic conditions ranging from almost perfect quiet to dense, complex sound environments? To address this question, we recorded extracellularly from the auditory thalamus of anaesthetised mice during presentations of spectrotemporally rich dynamic random chord (DRC) stimuli with varying spectral density (number of tone pips per octave). We quantified how the DRC-driven responses were modulated over time by decomposing the total response power (variance over time) into stimulus-dependent "signal power" and stimulus independent "noise power" (1). We found that the signal power decreased as stimulus density increased, while noise power slightly increased. This finding indicates that a sparse stimulus is capable of eliciting greater modulation of the neural response than a dense, more complex stimulus. We then fit the data using both linear spectrotemporal receptive field (STRF) models, and variants of multilinear context models (2) that capture nonlinear local interactions related to forward suppression and combination sensitivity. As the stimulus density increased, the fraction of signal power predicted by linear STRF models decreased, suggesting that encoding nonlinearities play a more significant role as the stimulus becomes denser and more complex. The predictive power of the multilinear context models was substantially higher than that of the linear STRF models at all stimulus densities (70-90% for the multilinear context models, 40-60% for linear STRF models), and the advantage did not change significantly as stimulus density increased. However, both the strength and latency of nonlinear delayed suppression captured by the context model decreased with increasing stimulus density. Our results show that auditory thalamic processing involves significant nonlinear contextual interactions in both sparse and dense acoustic environments; and may help to elucidate the role of nonlinear temporal integration in that processing.

References
Cognitive abilities such as language are supported by uniquely human adaptations that pose considerable challenges for understanding the evolutionary precursors to the relevant neural networks in the human brain. For instance, the ability to evaluate syntax—the grammatical relations between words in an utterance—is fundamental to human language but is difficult to study in nonhuman animals. Yet, a key component of this capacity is the ability to learn how expressions are appropriately sequenced. We propose the novel hypothesis that a precursor system for this core aspect of syntactic function exists in nonhuman primates, enabling them to evaluate whether sequences of auditory elements violate a previously learned set of rules, potentially part of a larger ‘grammar’ of legal sequences. To address this we aimed to: 1) determine if Rhesus macaques can implicitly learn rule-based sequences of auditory nonsense words (an auditory ‘artificial-language grammar’), and, 2) use functional magnetic-resonance imaging (fMRI) to study the functional anatomy of the brain regions that are involved in evaluating the grammaticality of the sequences. The behavioural experiments were conducted by first exposing the animals for ~45 minutes to exemplary sequences that follow the artificial-language rules. This was followed by a testing phase in which randomly sampled ‘correct’ and ‘violation’ sequences were presented to the monkeys, and we measured the frequency and duration of orienting responses toward the audio speaker presenting the test sequence. Testing conducted in the animals’ home colony (n = 13 macaques) based on 3 raters scoring filmed responses showed a group effect of longer look durations to the ‘violation’ sequences, suggesting that rule learning had occurred. Testing conducted in the laboratory with more objective eye-tracking based measurements confirmed the colony results and appears to be sensitive enough to be able to measure learning in individual animals. Two of these animals were then scanned with fMRI. Initial fMRI results show a set of frontal and temporal lobe brain regions that are sensitive to violations of the artificial grammar rules (fMRI activity contrast, ‘violation’ > ‘correct’ sequences). In summary, evaluating the ability of macaques to learn rule-based sequences has set the stage for fMRI to reveal potential evolutionary precursors to the brain regions that support syntax-related processing in humans. This may help to establish a non-human primate model system for understanding key functional aspects of the language-related network in the human brain.

Supported by: Wellcome Trust Project Grant (WT092606AIA) to CIP.
Inclusion or exclusion: encoding concurrent acoustic events in auditory cortex

Yi Zhou and Xiaoqin Wang
Laboratory of Auditory Neurophysiology, Dept of Biomedical Engineering, Johns Hopkins University, Baltimore, MD

Abstract:

In everyday listening, the detection of a target sound is often challenged by background noise. While noise masking has been ascribed to the “capture effect” of neural responses to noise at low signal-to-noise ratios, less is known about how auditory neurons encode concurrent acoustic events that are not sufficiently masked by each other. This study investigated tone-in-noise (T+N) detection by single neurons in the auditory cortex of awake marmoset monkeys. The experiment compared the level responses of neurons to T with and without N using both rate and temporal metrics. We found that a significant portion of cortical neurons exhibited non-monotonic rate-level characteristics to best-frequency (BF) tones – discharge rates are low at both low and high T levels, causing ambiguity in discerning the presence of a target based on the rate metric. On the other hand, we found that the temporal patterns in T+N responses could reliably signal the presence of either T or N or both, depending on the T level. The pattern was dominated by noise-like responses at low T levels and tone-like responses at high T levels, while T and N patterns coexisted at moderate T levels. Importantly, target- and noise-like responses extended to both driven and inhibitory epochs in neural activity. We argue that single neurons in auditory cortex coordinated T and N patterns in their responses through mechanisms of inclusion and exclusion, so that sensory information can be transmitted either together in the inclusion mode or alternately in the exclusion mode. These findings demonstrate the versatility of stimulus coding strategies used by single neurons and the availability of temporal pattern information in auditory cortex for encoding concurrent acoustic events in the environment.
How the auditory cortex drives decisions
Petr Znamenskiy, Anthony M. Zador
Cold Spring Harbor Laboratory

The mechanisms by which cortical sensory representations are transformed into voluntary actions remain largely unknown. We have therefore set out to assess the role of long-range projections from primary auditory cortex (ACx) in mediating sensory decisions. Here we report that the projection of the ACx to the auditory striatum drives auditory decisions in rats.

We first developed a novel auditory discrimination task designed to exploit the tonotopic organization of the ACx. We trained rats to discriminate low- and high-frequency “cloud-of-tones” stimuli in a two-alternative choice task. Each stimulus consisted of a sequence of short (30 ms) overlapping pure tones, distributed over a 3 octave range (5-40 kHz). Subjects were required to report whether tones of the low (5-10 kHz) or high (20-40 kHz) octave were overrepresented. Subjects’ performance varied smoothly with the fraction of low or high tones.

We hypothesized that the behavioral responses to the high- (or low-) frequency stimulus involved activation of the corresponding high- (or low-) frequency area of the tonotopic map, and that increasing neural activity in the high- (or low-) frequency region of the map would cause a corresponding decision bias. To test this hypothesis, we expressed ChR2 in ACx, and implanted arrays of optical fibers coupled to tetrodes, allowing concurrent photostimulation and recording of neural activity. As predicted, we found that ACx photostimulation during stimulus presentation biased the animals' responses. These results support a role for the primary ACx in mediating decisions in this task.

We next hypothesized that activity of ACx neurons projecting to the striatum would recapitulate this behavioral bias. In initial experiments, we found that the cortico-striatal projection is topographic, so that axons targeting a particular region of the striatum arise from a particular region within the tonotopic map in ACx, and individual sites within the striatum were tuned to sound frequency. We could therefore test this hypothesis by using optical fibers implanted in the auditory striatum to excite ChR2-expressing axons arising from neurons in ACx. Stimulation of these cortico-striatal fibers resulted in behavioral biases comparable to those obtained by stimulating the auditory neurons directly. These results indicate that the ACx projection to the striatum is sufficient to drive auditory decisions, and that the auditory striatum is engaged in the transformation of sound-evoked responses into action.