Disentangling sensory specialization from task specialization in lateral frontal cortex.

Abigail Noyce (1), Nishmar Cestero (2), Samantha Michalka (3), David Somers (1), Barbara Shinn-Cunningham (1)

(1) Boston University, (2) Columbia University, (3) Olin College of Engineering

Prior work has suggested that much of human lateral frontal cortex (LFC) is part of a 'multiple demand' network, supporting a wide range of cognitive tasks and processes (e.g. Duncan 2010; Fedorenko et al. 2013). In contrast to this multiple-demand view, several groups have recently reported that portions of LFC show a preference for processes associated with a particular sensory modality (Michalka et al. 2015; Braga et al. 2016; Mayer et al. 2016). Michalka et al. demonstrated that selective spatial attention to visual or auditory stimuli recruits specific structures within LFC. Two bilateral regions are biased for visual attention, superior and inferior precentral sulcus (sPCS and iPCS), and two bilateral regions are biased for auditory attention, transverse gyrus intersecting precentral sulcus (tgPCS) and caudal inferior frontal sulcus (cIFS). Here, we used fMRI to replicate this finding in a substantially different task paradigm and to examine the 'multiple demand' responsiveness of these regions. We observed that visual and auditory 2-back working memory tasks recruited interleaved visual- and auditory-biased structures that correspond to those reported by Michalka et al. These structures were mapped in each individual subject (n = 15). Within-subject reliability (n = 7) was high across the two tasks. We then tested the degree to which these sensory-biased structures demonstrated multiple demand behavior. First, we measured BOLD recruitment in the non-preferred memory task; visual-biased sPCS and iPCS showed significantly more recruitment during auditory WM than did tgPCS and cIFS during visual WM. Second, for each vertex within these structures we projected the activation in the two tasks into a 2D vector space; the multiple demand index derived from these values was significant for visual-biased but not for auditory-biased LFC structures. Both metrics indicate that the visual-biased LFC structures exhibit stronger multiple-demand responsiveness than do the auditory-biased structures, suggesting that the auditory processing network may be more specialized in its cognitive role. These results reconcile two competing theories of LFC’s contribution to human cognition by demonstrating the coexistence of sensory specialization and multiple demand behavior.
Data-driven segmentation of mouse auditory cortical fields based on mesoscale optical Ca2+ imaging

Sandra Romero (1), Ariel Hight (1), Daniel Polley (3), Jenifer Resnik (4)

(1) Eaton-Peabody Labs., Massachusetts Eye and Ear Infirmary, Boston, MA
(2) Dept. of Otolaryngology, Harvard Med. Sch., Boston

Microelectrode mapping studies have identified two tonotopically organized fields in the mouse auditory cortex (ACtx), the primary ACtx (A1) and the anterior auditory field (AAF), surrounded by a secondary auditory field (A2), a dorsoposterior field (DP) and the insular auditory field (IAF). Interest in cortical organization at “mesoscale” resolution (i.e., roughly 50 microns) has been revived thanks to optical imaging of genetically encoded calcium indicators (GECIs) that offer a higher signal-to-noise ratio than intrinsic signals, flavoproteins or bulk-loaded calcium indicators. Widefield epifluorescence imaging of GECIs in mouse ACtx has confirmed the basic organization published in earlier microelectrode mapping studies, with two unexplained discrepancies: First, a “dead” zone between A1 and AAF that does not respond to tones (but may instead respond to more complex sounds such as FM sweeps). Second, high-frequency regions lateral to the principal tonotopic vectors that link A1 and AAF. These two differences, in combination, have inspired a variety of new naming conventions and field demarcations for the mouse ACtx. The principal motivation for this study was to perform mesoscale optical imaging of GECIs in the mouse ACtx and develop objective, data-driven approaches to parcel the various fields. Widefield GCaMP6s imaging through chronically implanted cranial windows in awake mice confirmed the basic spatial layout of A1, AAF, A2, DP and IAF. We also observed the existence of high-frequency regions lateral to A1 and AAF as well as a region at the center of the ACtx that was weakly activated by pure tones. A closer examination revealed that this central region was tone-responsive, albeit with higher thresholds and lower response amplitudes. A pixel-based thresholding approach confirmed that best frequencies in this region formed a single tonotopic reversal between A1 and AAF. With this analytic adjustment in place, it was evident that tonotopic gradients within A1 and AAF did not bifurcate into two parallel lateral and medial fields, but instead formed a continuous tonotopically organized field. We developed an objective analysis approach to establish boundaries between fields based on points of tonotopic reversal or abrupt drops in signal amplitude. This analysis approach reveals strong tonotopic organization in A1 and AAF that “fans” out from low-frequency regions and collides at a single medial-lateral boundary, consistent with other species. Repeated imaging in the same mouse suggested that organization was reliable, with only minimal day-to-day variations, at least at this relatively coarse spatial scale.
3 Neural oscillatory activities for processing dynamic auditory information: from intrapersonal to interpersonal entrainment

Andrew Chang (1), Dan J. Bosnyak (1), Laurel J. Trainor (1)

(1) McMaster University

Humans must process highly dynamic information, such as in speech and music, in order to decode meaning and react in real time. Prior studies showed that sensory systems proactively predict upcoming information to optimize perceptual and sensorimotor processing. Predictive processing may be represented by neuronal oscillations. In particular, the power of induced beta oscillations (15 – 25 Hz) in auditory and motor areas entrains to the rate of a presented isochronous tone sequence (e.g. Fujioka et al., 2012, JN), with reductions in beta power following sound event onsets, and with the slope of the rebound predicting the onset of the next sound event. However, it remains unclear whether beta entrainment facilitates perceptual performance intrapersonally, and whether complex interpersonal cooperative activities such as group musical performance can be represented by oscillatory entrainment between individuals. In the first study, we investigated how oscillatory neural activity intrapersonally entrains to an external auditory sequence, and how it relates to perceptual processing. By combining EEG and psychophysical techniques, we presented identical tones in rhythmic versus arrhythmic sequences; occasionally, one tone was replaced by a target tone with modified pitch, and participants were instructed to discriminate whether target pitches (presented louder) were higher or lower than the standard pitch. We investigated how pre-target oscillatory neural activity predictively determines post-target perceptual performance. The results showed 1) pitch discrimination sensitivity was higher when target tones were embedded in rhythmic than arrhythmic sequences, 2) pre-target beta power entrained to the temporal regularity of the sequence, and 3) trial-by-trial analyses found that the size of pre-target beta entrainment positively predicted pitch discriminative sensitivity. These results indicate that beta entrainment improves perceptual processing for dynamic auditory information. In the second study, we investigated dynamic sensorimotor processing in a real-world interpersonal interaction, using professional string quartets as a model for interpersonal coordination. We experimentally manipulated leadership roles during performances, and employed Granger causality to investigate the directional coupling among musicians. Analyses on body movements revealed that directional coupling followed assigned leadership (Chang et al., in press, PNAS), as predicted. Ongoing analyses are examining EEG dynamics, especially in beta band, to investigate oscillatory synchronization and information flow between brains."
4 Auditory and visual sequence learning in humans and monkeys

Alice E. Milne (1,2), Su Li (3), Christopher I. Petkov (1,2) & Benjamin Wilson (1,2)

1 Institute of Neuroscience, Newcastle University, UK
2 Centre for Behaviour and Evolution, Newcastle University, UK
3 Cambridge University, UK

Language flexibly supports the human ability to communicate using different sensory modalities. Moreover, syntactic operations engage certain overlapping regions in the human brain during, for instance, reading in the visual modality and listening to language in the auditory domain. Although it has been argued that nonhuman primate communication is inherently multisensory, there are few direct comparisons of human and nonhuman primate abilities across sensory modalities. Artificial Grammar Learning (AGL) tasks can emulate ordering relationships between words in a sentence, however comparative AGL work has primarily investigated sequence learning in a single modality. We used an AGL paradigm to evaluate how humans and macaque monkeys respond to identically structured sequences of non-linguistic auditory or visual stimuli. In both modalities, the two species gave remarkably similar response patterns in the visual and auditory domains, indicating that the sequences are processed comparably across the sensory modalities. We next conducted an fMRI experiment using the same auditory and visual sequences in humans, and an equivalent macaque fMRI experiment is underway. Multi-voxel pattern analysis of the human fMRI data using Representational Similarity Analyses (RSA) investigated brain responses in sensory cortices and other brain regions. The results show that, primary sensory cortices respond to sequencing relationships within their respective sensory modalities. Brodmann areas 44/45 and the frontal operculum in frontal cortex respond to aspects of the sequence ordering relationships in both modalities. These observations identify the domain-specific sequencing operations in the respective sensory processing streams, and they highlight inferior frontal cortex as a domain-general substrate for sequence processing. The results also provide initial evidence that human sequence learning abilities stem from an evolutionarily conserved capacity for multisensory sequence processing.

Supported by Wellcome Trust
Auditory neurofeedback training of selective attention and speech recognition in noise

Subong Kim (1), Adam T. Schwalje (2), Caroline Emory (1), Inyong Choi (1)

(1) University of Iowa
(2) University of Iowa Hospitals and Clinics

Speech recognition in noise is a hard task that involves multiple stages of cortical processing including acoustic feature extraction, auditory grouping, working memory and attention. Among those processes, we aimed to investigate whether auditory selective attention ability affect speech-in-noise understanding performance, and whether the neurofeedback training of selective attention can improve speech understanding in noise. Selective attention enhances the strength of auditory N1 response to attended sounds while suppresses the responses to ignored sounds, which forms an evidence of sensory gain control theory. We hypothesized that the N1-amplitude guided neurofeedback can strengthen the cortical sensory gain control, which in turn will improve selective attention performance and may result in better speech-in-noise understanding. With a single-blinded, between-subjects design including a control (placebo) group, subjects were asked to attend to one of two simultaneous auditory streams of single-word repetitions (e.g., five “up”s and four “down”s from left and right loudspeakers). For the participants assigned to the experimental group, single-trial EEG responses were classified using template-matching method based on pre-defined ideal EEG patterns (i.e. grand-average EEG responses elicited by either “up” or “down” stream). A visual feedback was provided after each trial to demonstrate whether the EEG signal was correctly classified or not. This attention-driven brain-computer interface (BCI) worked well and nearly all subjects inspected that their selective attention was decoded by EEG with real time. The experimental group participants with 4 weeks of this neurofeedback training improved their ability to understand speech-in-noise, while the placebo group participants who only received behavioral feedback did not show consistent improvement. To the best of our knowledge, this is the first report of selective-attention training enhancing speech-in-noise understanding ability.
6 Auditory attention and predictive processing co-modulate speech comprehension in middle-aged adults

Sarah Tune (1), Malte Wöstmann (1), Jonas Obleser (1)

(1) University of Lübeck

In real-life communication, speech comprehension requires the dynamic engagement of a complex set of perceptual, executive control and prediction processes. This challenge becomes exacerbated by the gradual declines in sensory acuity and cognitive functioning that are typically associated with healthy aging starting in middle adulthood. Here, we present results from a study focused on healthy middle-aged human adults (40–70y) who performed a novel dichotic listening task. The paradigm called for adaptive control of cognitive strategies by varying the degree to which auditory spatial attention and predictive processing support comprehension.

Participants were presented with two competing, dichotically presented speech streams uttered by the same female speaker. Participants were probed on the last word in one of the two streams. Crucially, auditory presentation was preceded by two visual cues. First, a spatial-attention cue either indicated the to-be-probed side, thus invoking selective attention, or it did not provide any information about the to-be-probed side, thus invoking divided attention. The second cue specified a general or a specific semantic category for the target word (and was valid for both utterances). This semantic cue therefore facilitated semantic and sensory prediction of the upcoming input.

Behavioral results (n=29) show a general increase in performance for informative compared to uninformative cues. Participants responded faster in selective attention trials and following a specific semantic cue. Accuracy was co-modulated by the joint effect of both cues, as reflected by a benefit from specific (vs. general) semantic cues but only under selective auditory attention. Moreover, reliance on the spatial-attention cue varied with age: Older adults performed better under selective attention and worse under divided attention than younger adults.

Analysis of electroencephalography (EEG) data (n=16) revealed a lateralization of 8–12Hz alpha power during spatial attention cue presentation, but also and even more pronounced during the dichotic speech streams in selective-attention but not in divided-attention trials. Specificity of the semantic cue on the other hand modulated oscillatory power in the beta frequency band (15–30Hz), with a decrease in power for specific cues.

In sum, our results provide evidence for the interplay of attentional control and predictive processes in difficult listening situations. Crucially, providing two distinct types of cues prompted changes in behavioural performance correlated with qualitatively different neural signatures, and highlights changes in cognitive strategies with age."
Mice emit ultrasonic vocalizations (USVs) which vary in spectrotemporal parameters (e.g., frequency, amplitude, and duration) in various social situations. USVs are often assumed to possess speech-like characteristics, although it has not yet been completely established that mice are using USVs for communication purposes. Previous studies have shown changes in auditory cortex activity in maternal females to pup calls, but it is currently unknown how previous social experience with other mice throughout development affects the perception of adult vocalizations. To test the effects of socialization, we used an operant conditioning task to determine if discrimination of USVs was negatively impacted by chronic social isolation compared to mice that were group housed throughout their lifespan. Mice discriminated between eighteen USVs of three different categories. Mice that had been socially isolated since weaning showed differences in discrimination of USVs relative to socially experienced mice. Additionally, socially isolated mice initially required more training and testing, and more trials to complete the task than socially experienced mice, with improvements in performance associated with experience with the experimental stimuli. These results indicate that active experience with USVs, through either social interactions or from repeated exposure during the operant task, can affect how the mice perceive vocalizations. Results suggest that these vocalizations could have context-specific meaning that is learned through hearing USVs within the appropriate social context, which gives socialized mice an advantage early in testing, but experience with USVs throughout testing also leads to a change in the perception of calls. Supported by NIH DC012302.
A dynamic neural code may underlie multisensory integration and segregation in the primate superior colliculus

Jeff Mohl (1), Surya Tokdar (1), Jennifer Groh (1)

(1) Duke University

At the neural level, multisensory integration has typically been characterized in relation to either enhancement over the strongest component unisensory response (including linear summation of both components), or as an averaging computation between the unisensory responses (sometimes referred to suppression, as it suppresses activity below the best unisensory response). This averaging effect is particularly pronounced in conditions of cue conflict, such as mismatched visual and vestibular cues or misaligned lights and sounds. However, conflicting cues often result in separate percepts, indicating each is simultaneously and separately represented in the brain, apparently at odds with pure averaging. One alternative possibility is that only one stimulus is encoded by each neuron or population at a given time, and that switching between representations results in apparent averaging when activity is pooled across time and multiple trials.

To probe this question, we presented monkeys with either visual only, auditory only, or combined visual and auditory stimuli at varying locations in space, while we recorded the activity of superior colliculus neurons. We analyzed the data using a novel statistical assessment capable of characterizing neural responses at the single trial level (Caruso et al., bioRxiv 2017). We found that responses on combined modality trials often showed activity fluctuations across trials consistent with time-division multiplexing of the individual modalities. This suggests the existence of a dynamic neural code that retains information about both components of multimodal stimuli."
9 Distinct maturational trajectory of temporal processing in thalamocortical recipient layers 4 and 5

Fenghua Xie(1), Ling You(1), Dongqin Cai(1), Miaomiao Liu(1), Yin Yue(1), Yiwei Wang(1), Kexin Yuan(1)

(1) Tsinghua University

As in other primary sensory cortices, both layer 4 (L4) and layer 5 (L5) of the primary auditory cortex (A1) receive substantial thalamic inputs. Temporal processing in L4 of both developing and adult A1 has been extensively studied. However, conclusions about temporal response properties in L5 compared to L4 are still controversial. Furthermore, the maturational process of temporal processing in L5 remains unknown. Using loose-patch recordings in-vivo, we found that the temporal response resolution of L5 neurons in the adult rat A1 is not significantly different from that of L4 neurons. However, very interestingly, L5 neurons exhibit superior stimulus-following ability immediately after hearing onset, in contrast to the poor temporal responses of L4 neurons. In fact, no significance difference in temporal processing was observed between L5 neurons in the developing and adult A1. In-vivo whole-cell voltage-clamp recordings showed that L4 and L5 neurons in the developing A1 are not significantly different from each other in terms of both resting membrane potential and the adaptation of both excitatory and inhibitory inputs, although L5 neurons are more depolarized and the adaptation of inhibition is stronger in the adult A1. However, the duration of inhibitory input to L5 neurons is remarkably shorter than that to L4 neurons in the developing A1, while they are similarly short in adults. Moreover, the difference in inhibitory duration between L5 of developing and adult A1 is only a fraction of that between L4 of the two age groups. Repetitive stimulation does not evoke summation of inhibition in L5, keeping the classical temporal sequence between excitation and inhibition. These results have several implications. First, they further confirmed that the duration of inhibition plays a critical role in setting the pace for the maturation of cortical temporal processing. Second, they suggest that the inhibitory circuits in L4 and L5 are distinct, at least partly underlying the difference in the developmental trajectory of functional properties. Last but not least, they support the notion that although both L4 and L5 are thalamocortical recipient layers, they may represent two separate systems subserving distinct functions.
The auditory system of two pinniped species

Juliane Krueger (1), Emily C. Turner (1), Eva K. Sawyer (1), Jon H. Kaas (1);

(1) Department of Psychology, Vanderbilt Univ., Nashville, TN

Pinnipeds (sea lions, seals, and walruses) are carnivorous mammals who have adapted to a life on land and at sea. Very little is known about their central nervous system but recent studies have begun to characterize their sensory systems. Their ape-size brains are marked by an extensive arrangement of sulci and gyri and thus far appear to be similar in their organization to other carnivores such as cats. Here, we extend past anatomical studies from our lab of the Northern sea elephant (Mirounga angustirostris) and the California sea lion (Zalophus californianus) investigating their somatosensory (Sawyer et al, 2016) and their visual (Turner et al, 2017) networks to include the auditory system. Utilizing Nissl substance, cytochrome oxidase, and vesicular glutamate transporters (VGluT 1 and 2), we identified the auditory brainstem, midbrain, and thalamic nuclei in addition to primary auditory cortex. We found that the cytoarchitectural organization of the cochlear nucleus (CN), the superior olivary complex (SOC), the inferior colliculus (IC), the medial geniculate nucleus (MGN), and the core auditory region (A1) appeared to be similar between the elephant seal (n=2) and sea lion (n=2), with the former featuring slightly larger nuclei across the auditory system. Cytochrome oxidase stained sections of the elephant seal for example revealed a dorsal and ventral division of the CN, at least two subnuclei (the medial and the lateral divisions respectively) of the SOC, a clear division between the dense central nucleus and the dorsal cortex of the IC, the basic MGN subdivisions, and a strong layer IV band in auditory cortex. Similar observations were made with Nissl and VGluT2 in elephant seal and appeared to also hold true in the sea lion brain. Several auditory regions of these two pinniped species exhibited strong similarities with the auditory system of cats making cats an ideal reference for nuclei identification. Furthermore, some architectural similarities were also observed in ferrets and dogs. Thus, these results may provide additional insight in the evolution of large brains supporting a conservation of a set of auditory nuclei across species.
Live music increases intersubject synchronization of audience members' brain rhythms

Molly J. Henry (1), Daniel J. Cameron (1), Dana Swarbick (2), Dan Bosnyak (2), Laurel Trainor (2), & Jessica A. Grahn (1)

(1) University of Western Ontario
(2) McMaster University

Attending concerts is enjoyable for a number of reasons: watching performers make live music affords a qualitatively different experience than listening to a recording. Moreover, an important contributor to the enjoyment of a concert—at least anecdotally—is forming a bond with others who are enjoying the same musical experience. The current study considered the possibility that a live musical experience, i.e., the presence of live performers as well as an audience, might change the way brain rhythms synchronize across audience members, thereby changing audience members' musical and affiliative experiences. We collected electroencephalography (EEG) data in three different social contexts. First, EEG was measured simultaneously from 20 audience members (in a larger crowd of approximately 80 people) while they observed a live musical performance. Second, EEG was measured simultaneously from 20 audience members (in a larger crowd of approximately 80 people) while they watched the recording of the first concert on a large movie screen and with audio identical to the live concert. Finally, EEG was measured from 20 participants in small groups of 2 participants seated apart (tested in 10 separate sessions) while they observed the recorded musical performance. Thus, we manipulated the presence of the performers while keeping audience context fixed, and we manipulated the presence of other audience members while keeping the recorded performance fixed. We analyzed the data in terms of intersubject synchronization (ISS), which quantifies the degree to which brain rhythms are synchronized across groups of individuals. ISS was calculated for individual frequencies ranging between 0.1 Hz (‘infra-slow’ oscillations) to 60 Hz (gamma-band oscillations) for each social context condition. We observed differences in the delta (2-4.5 Hz) and low-beta (13-16 Hz) bands depending on the presence of the performers—that is, audience members’ brain waves were more synchronized with each other when the performers were present. The delta band corresponds roughly to the range of rates in which a musical beat would be felt, and beta band brain rhythms, in addition to having strong associations with movement and motor system, critically have been linked to timing and temporal prediction in rhythmic sequences. ISS was similar across conditions that involved watching a recorded performance, whether other audience members were present or not. Thus, the presence of live performers at a concert leads to increased synchronization of audience members’ brain rhythms selectively in frequency bands that are associated with feeling and moving along with a musical beat.
The inferior colliculus (IC) is a major hub of auditory processing. After different aspects of an acoustic signal are processed in the auditory brainstem nuclei, this information converges in the central nucleus of the IC (ICC) onto one tonotopic map. According to the synaptic domain theory, the inputs from different auditory brainstem nuclei to the ICC cluster in specific sub-regions and, thus, form functional zones superimposed on the single tonotopic map. These sub-regions may send this information via the thalamus to distinct areas of the core auditory cortex with different functions. One approach to the detailed study of function and anatomy in the ICC sub-regions is to genetically manipulate the IC during development.

In the present study, we used a mouse model with an oversized IC. Developed by Dee et al. (2016), MEK1 (also referred to as mitogen-activated protein 2 kinase 1, MAP2K1) is overexpressed in this mouse in the tectal stem cell zone from which the IC originates. Consequently, the stem cells remain in the proliferation stage longer, thus increasing the number of stem cells, but also delaying neurogenesis. This results in more, but later developed IC neurons. The gross anatomy shows that in all MEK1 mice the IC is massively enlarged compared to littermate controls.

We tested the hearing thresholds of MEK1 mice using click-evoked auditory brainstem response (ABR) and amplitude-modulated frequency following response (AMFR) measurements. To obtain the AMFR audiogram, we used narrow-band noise (0.3 octave) centered at frequencies 2-40 kHz and a modulation frequency of 42.9 Hz. Our preliminary electrophysiological evaluation of the MEK1 mice showed a diverse phenotype. In comparison to the audiograms of littermate controls, some MEK1 mice had an almost normal audiogram, while others showed elevated thresholds. We also measured the amplitude growth in the AMFR signal in response to increasing sound level intensity. In MEK1 mice there was a reduced growth function. Interestingly, the hearing threshold did not predict the amplitude growth function or vice-versa. We also measured the peak synchrony of the AMFR signal and found that synchrony was slightly degraded in most MEK1 mice compared to littermates. In summary, our preliminary observations of a mouse with a massively enlarged IC suggest that this structural change may result in more than one type of alteration in the circuitry of the auditory midbrain and more than one hearing phenotype. Bigger may not be better.

Supported by NIH R21DC013822 and UConn Spring 2017 Health Research Program (N.M.)
Transforming continuous temporal cues to a categorical spatial code in human speech cortex

Neal P. Fox (1), Matthias J. Sjerps (2), Matthew K. Leonard (1), Edward F. Chang (1)

(1) University of California, San Francisco
(2) University of California, Berkeley; Radbound University

During speech perception, listeners extract acoustic cues from a continuous sensory signal to map it onto phonetic categories. Many such cues are encoded within the fine temporal structure of speech. For example, voice-onset time (VOT), the interval between a stop consonant’s release and the onset of voicing, distinguishes voiced (e.g., /b/, short VOT) from voiceless (e.g., /p/, long VOT) stops in English. The neurophysiological mechanisms that allow listeners to distinguish sounds that differ in temporal cues remain unclear. We recorded neural activity directly from the cortex of 9 human subjects while they listened to and categorized syllables along a VOT continuum from /ba/ (0ms VOT) to /pa/ (50ms VOT). We found that spatially distinct neural populations respond preferentially to one category (either /b/ or /p/). In both populations, responses are sensitive to VOT differences within the preferred, but not the non-preferred, category. This graded VOT encoding rapidly evolves to reflect the behavioral response function, showing that categorical perception of VOT emerges across time in auditory cortex. To probe what computations might give rise to these response properties, we implemented a neural network model that simulates neuronal populations as leaky integrators tuned to detect either coincident or temporally-lagged burst and voicing cues. The same temporal dynamics and encoding patterns observed in real neural data emerged in the model, suggesting that local tuning for distinct spectral cues at precise lags may underlie temporal cue integration in auditory cortex. Our results provide direct evidence that continuous temporal information is transformed into a categorical spatial code by discrete, phonetically-tuned neural populations in human auditory cortex.
We experience space as a unitary/coherent whole, regardless of which senses are involved in detecting the stimuli that make up the scene. However, spatial information acquired through different senses is known to be represented in the primate brain in different formats: visual locations are encoded in maps, while auditory locations are encoded in meters (or rate) codes.

The discrepancy in coding format across sensory modalities raises the question of how and where such differences get resolved. Here, we quantitatively compare the visual and auditory representations in three brain areas that contribute to guiding eye movements to visual and auditory targets: the SC, the frontal eye field (FEF), and the lateral and medial intraparietal cortex (LIP/MIP). We investigated single cell and population coding of visual and auditory locations during a saccade task using identical methods in all areas.

We found that single cell tuning curves for visual and auditory stimuli were broad and more similar to each other in the cortical areas compared to the SC. However, the degree of similarity between individual visual and auditory tuning curves was not reflected in the pattern of correlation between neural population vectors and was not predictive of the population encoding of space when evaluated with principle components analysis. Whereas the SC and FEF encoded space almost independently of modality at the population level (with a relative compression of the auditory space), the LIP/MIP encoded visual and auditory space in two largely distinct populations, compressing and possibly distorting the auditory space with respect to the visual space.

These results suggest an alternative method of evaluating coding format beyond maps and meters and their associated closed vs. open fields. The pattern of results across areas suggests different fidelities of spatial representations, possibly subserving different roles, such as multisensory integration vs. precise movement execution"
Transcranial 10-Hz stimulation but also eye closure modulate auditory attention

Malte Wöstmann (1), Lea-Maria Schmitt (1), Johannes Vosskuhl (2), Christoph S. Herrmann (2), Jonas Obleser (1)

(1) Department of Psychology, University of Luebeck, Germany
(2) Experimental Psychology Lab, Center for Excellence Hearing4all, European Medical School, University of Oldenburg, Germany

When humans focus attention to auditory events, neural alpha oscillations (~10 Hz) in the Magneto-/Electroencephalogram (M/EEG) increase in power. Here we test whether experimentally induced increases in alpha power modulate auditory attention. In two studies, healthy human participants attended to spoken target digits against distractors. Alpha power was increased exogenously through transcranial alternating current stimulation (tACS), or endogenously through eye closure.

In study I (n = 20), participants were cued to attend to a stream of four spoken digits presented to one ear, while ignoring a distracting (same-talker) stream of digits presented to the other ear. Previous M/EEG studies have shown that such dichotic tasks increase alpha power in auditory and parietal cortex ipsilateral to the focus of attention. To manipulate this alpha lateralization, we applied continuous 10-Hz tACS to temporal and parietal scalp regions in the left hemisphere (1 milliamp; stimulation sites FC5 and TP7). To control for the effect of stimulation frequency, each participant also received sham and gamma-tACS (47.1 Hz). Compared to sham, left-hemisphere alpha-tACS enhanced the recall of target digits in ‘attend-left’ versus ‘attend-right’ trials, while the opposite was found for gamma-tACS. This suggests that an exogenous increase in lateralized alpha power relatively suppresses auditory spatial attention to the side opposite to stimulation.

In study II (n = 22), we sought to invoke an endogenous increase of alpha power instead while presenting participants with two alternating (different-talker) streams of five spoken digits. On each trial, participants were instructed to attend to one stream and to ignore the other. In blocks where they closed their eyes (in a dark room; compared to keeping their eyes open), participants induced a baseline increase in parieto-occipital EEG alpha power. During a trial, baseline-corrected alpha power fluctuated rhythmically, with alpha peaks preceding onsets of attended digits by ~100 ms. This attentional modulation of alpha power strongly increased with closed compared to open eyes, demonstrating that eye closure boosts the neural difference in auditory attending versus ignoring. However, eye closure did not enhance participants’ ability to afterwards tell attended from ignored digits, which contradicts the widely held belief that eye closure enhances the behavioral outcome of attentive listening.

In sum, the observed impact of eye closure and alpha-tACS on neural alpha dynamics and behavioral corollaries suggest that alpha power is more than a mere
epiphenomenon but neurally and behaviorally relevant to auditory attention."
Objectives: We investigated whether we can classify auditory EEG while avoiding imposing a model how speech is processed in the brain. Currently, measuring speech entrainment is mostly done by taking stimulus features into account, such as constructing a linear decoder which estimates the speech envelope. This method imposes restrictions, as it assumes stimulus features and a model of the brain. We present a model-less approach for EEG classification based on artificial neural networks (ANN).

Methods: Nine normal-hearing subjects listened attentively to repetitions of the Flemish Matrix speech material while their EEG was recorded. These Flemish Matrix sentences are 5 word sentences of approximately 2 seconds, for each word there are 10 possibilities. Five subjects also listened to these sentences combined with speech weighted noise at different signal-to-noise ratios (-12 dB SNR, -8 dB SNR and +4 dB SNR). We trained an ANN to classify which sentence was presented given the EEG signal and compared the results with a method using the speech envelope, which is used in a wide variety of research concerning coding of speech in the brain.

Results: The ANN approach yielded significantly higher classification performance than the speech entrainment approach, which uses the envelope of the stimulus. A subject-specific ANN and an ANN trained on all subjects were both able to classify the presented sentence. Finally, we found that the accuracy of the classification increased with increasing SNR.

Conclusion: Using an ANN we were able to estimate which sentence was presented given the EEG signal. With a subject-specific ANN accuracies dropped as the SNR of the stimulus decreased. This promising model-less approach can be interesting for future research in neural encoding of speech.

Acknowledgements: Research funded by a PhD grant of the Research Foundation Flanders (FWO). This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 637424).
Sensory signals originating from the same event, such as the voice and mouth movements of a speaker, are often temporally correlated. It is hypothesized that the brain evaluates these correlations to facilitate feature integration and binding within and across sensory modalities. Previous studies have shown that correlation between unisensory signals facilitates several multisensory behaviors, whereas uncorrelated signals do not. In the current study, we sought to further illuminate the nature of this relationship, hypothesizing that multisensory behavior will vary with the strength of correlation. To this end, we presented participants with sinusoidal amplitude modulated auditory and visual stimuli at modulation depth thresholds. Participants reported the presence or absence of modulation in either stimulus. On a trial, modulation was present in auditory, visual, both, or neither modality (20% of the total trials, catch trials). On audiovisual trials, visual modulation frequency and phase were constant while auditory modulation frequency and phase were varied to generate stimulus pairs with a range of temporal correlations. Accuracy and reaction time data were fit to a drift diffusion model in which non-decision time and drift rate could vary across conditions. The pattern of discriminability and drift rate across conditions was very similar to that of stimulus correlation, but with an apparent phase shift that reflected unique temporal processing of the unisensory stimuli across participants. After accounting for this phase shift, drift rate varied with stimulus correlation in every participant, suggesting that stronger stimulus correlations provide stronger sensory evidence. In two participants, non-decision time decreased with increased stimulus correlation indicating faster encoding of positively correlated stimuli. These results indicate that the degree of stimulus correlation strongly impacts multisensory perception. Further, they suggest that the process of binding could be stochastically dependent on correlation such that signals are more likely to be bound as their correlation approaches 1.
Cortical processing of spectrally degraded speech revealed by intracranial electrophysiology

Mitchell Steinschneider(1)*, Kirill V. Nourski(2)*, Ariane E. Rhone*(2), Hiroto Kawasaki(2), Matthew A. Howard III(2)

*These authors have contributed equally to this work

(1) Albert Einstein College of Medicine
(2) University of Iowa

There is considerable variability in speech perception outcomes following cochlear implantation. Auditory cortical function and plasticity are hypothesized to account for much of this variability. Understanding the cortical processing of spectrally degraded speech can help define the roles of different auditory cortical areas in speech perception and contribute to improvements in CI design. A noise vocoder serves as an acoustic model of cochlear implant (CI) sound processing, allowing for investigation of responses to spectrally degraded speech in normal-hearing listeners. This study investigated electrocorticographic (ECoG) responses within auditory cortex elicited by noise-vocoded speech stimuli and related these responses to behavioral performance in a 2-alternative forced choice (2AFC) identification task.

Subjects were neurosurgical patients undergoing chronic invasive monitoring for medically refractory epilepsy. Stimuli were utterances /aba/ and /ada/, spoken by a male talker. The stimuli were spectrally degraded using a noise vocoder (1-4 bands). ECoG data were recorded during performance of the 2AFC task and were obtained simultaneously from Heschl’s gyrus (HG) and superior temporal gyrus (STG) using multicontact depth and subdural grid electrodes, respectively. Event-related band power was examined within the high gamma (70-150 Hz) frequency range.

All subjects performed at chance-level identification accuracy with speech degraded to 1 or 2 spectral bands, and performed at or near ceiling with non-degraded stimuli. Subjects exhibited a wide variability in accuracy at 3- and 4-band conditions. High gamma responses in posteromedial HG (auditory core cortex) were similar in magnitude for vocoded and natural stimuli and did not reflect behavioral performance. Responses from anterolateral HG were generally weaker and exhibited selectivity for natural stimuli. In contrast, responses from lateral STG to vocoded speech were larger in subjects who performed better in the task. In these subjects, responses increased in parallel with behavioral performance.

Findings highlight differences in representation of spectrally degraded speech across core and non-core cortical areas and their relationship to perception. These findings support earlier non-invasive results that demonstrate the relationship between robust non-core auditory responses and speech intelligibility. By modeling the patterns of cortical activity elicited by CI stimulation, intracranial data offer the opportunity to clarify
cortical processing of degraded speech and, in future studies, the improvements that occur in CI users with rehabilitation and experience.
Neural source dynamics of brain responses to continuous speech: from acoustics to comprehension

Christian Brodbeck (1), Alessandro Presacco (2), Jonathan Z. Simon (1)

(1) University of Maryland, College Park
(2) University of California, Irvine

Natural speech processing is inherently difficult to study with neuroimaging methods because of the temporal and acoustic irregularity of continuous speech; much of what we know about the neural basis of language comprehension is based on experimental designs that sacrifice naturalness of the stimuli, leaving open the question of whether neural processing differs in more realistic settings. To address this issue, we combined source estimation of magnetoencephalography (MEG) data with reverse correlation to predict the neural response to different properties of continuous speech in time as well as by anatomical location. We first computed distributed minimum norm current estimates for the MEG responses of young adults listening to segments of an audiobook, and then used reverse correlation based on the boosting algorithm to compute response functions for different predictors at each virtual current source. Response functions were statistically assessed with permutation tests. Results show neural processing of multiple levels of the speech stimulus: An early (~50 ms) response to the acoustic envelope of the speech stimulus separates into an earlier (~35 ms) response in auditory cortex, and a later (~50 ms) more dorsal response, consistent with the mouth area of somatosensory cortex, suggesting fast mapping of acoustic information to somatosensory/motor representations. A later response to the acoustic envelope (~100 ms) localized predominantly in right auditory cortex, while a response associated with word-frequency was localized to left auditory cortex, suggesting differential processing of acoustic and lexical information in the two hemispheres. Finally, a response to words that enable semantic composition is seen in higher level language areas, anterior temporal lobe and inferior frontal gyrus, demonstrating that even comprehension-related responses can be localized with reverse correlation. Our results indicate that MEG responses to continuous speech are rich in dynamic information that can be spatially reconstructed. This extends the set of possible hypotheses about the neural basis of speech comprehension that can be tested in the natural setting of listening to continuous speech.
Cortical responses in human superior temporal gyrus that differentiate intonation contours in speech are a response to pitch, not fundamental frequency

Claire Tang (1), Liberty S. Hamilton (1, 2), Edward F. Chang (1)

(1) University of California, San Francisco
(2) University of Texas, Austin

Speech intonation is an important component of prosody in all spoken languages. It conveys sentence-level linguistic meaning, such as sentence type (statement vs. question) or focus (which word was emphasized). In our previous work using electrocorticography (ECoG), we showed that specific neural populations in the human superior temporal gyrus (STG) respond differentially to speech stimuli with distinct intonation contours, realized through digital manipulation of the fundamental frequency (f0). Because those stimuli were synthesized to control for intensity and duration of syllables, the only acoustic difference across intonation conditions was f0 over time. Although the f0 of a sound generally determines its perceived pitch, sounds with acoustic energy at the fundamental removed, called 'missing fundamental' stimuli, can also produce a pitch percept. To determine whether STG responses that differentiated intonation contours in speech are a response to pitch or to f0, we created a set of non-speech stimuli that preserved each intonational pitch contour but did not contain f0 and played them to participants while we recorded cortical activity using ECoG. To test whether neural responses to pitch contours in the missing f0 stimuli were similar to neural responses to f0 contours in speech, we used linear discriminant analysis to fit a model predicting the intonation contour from neural responses to speech. We then tested this model on the missing f0 data to see whether the model could discriminate between responses to the missing f0 stimuli. We found that in almost all electrodes (47/49 electrodes from 3 participants), neural activity patterns to the different intonation contours were the same between speech and missing f0 stimuli. This result indicates that the cortical activity in neural populations that differentiate intonation contours in speech can be explained as a response to the perceptual attribute of 'pitch', rather than energy at the fundamental frequency.
Primate behavioral and functional-imaging model for auditory figure-ground segregation

Felix Schneider (1), Pradeep Dheerendra (1), Fabien Balezeau (1), Alexander Thiele (1), Timothy D. Griffiths (1)

(1) Institute of Neuroscience, Newcastle University, Newcastle Upon Tyne, United Kingdom

Segregating sounds in a noisy environment is a fundamental aspect of scene analysis. Inability to detect figures from noisy background is a ubiquitous problem in both cochlear hearing loss and in disorders of central sound processing. In normal human listeners, emerging evidence suggests that auditory objects are detected with remarkable sensitivity and robustness based on a mechanism that detects temporal coherence in different frequency bands (Teki et al. 2013). Human imaging studies demonstrate a system including auditory cortex in the superior temporal sulcus (STS) in non-core homologues and in the intraparietal sulcus (IPS) during preattentive, stimulus-driven figure-ground decomposition of stimuli (Teki et al. 2016; Teki et al. 2011). We have developed a primate model in a species in which the system organisation can be compared to humans more easily than in other mammals (Baumann et al. 2013). The eventual aim is to achieve an understanding of figure-ground analysis at the neuronal level based on systematic recordings that are not possible in humans. We present data from behavioural experiments to measure the ability of rhesus macaques (n = 2) to detect target sounds in a noisy background using stimuli and parameters determined by the previous human work (Teki et al. 2013). A Go/No-Go task with bar release measured the detection of figures based on keeping 10 frequencies constant in a random chord sequence where each chord had duration of 50ms. Parallel functional imaging using fMRI on 3 subjects demonstrated the network for detection of similar 10-component figures presented over 40 chords with a duration of 50ms. As in previous human fMRI (Teki et al. 2011) an irrelevant task was carried out with the aim of demonstrating preattentive, stimulus-driven figure-ground detection. Sparse imaging was used. A contrast between scans after presentation of figure plus ground and control trials after ground only demonstrated activity in parabelt homologues in anterior superior temporal gyrus (STG). In one subject we demonstrated significant activity in posterior STS.

References:


Teki, S. et al. (2013): Segregation of complex acoustic scenes based on temporal coherence. eLIFE 2013:2
Teki, S. et al. (2011): Brain Basis for Auditory Stimulus-Driven Figure-Ground Segregation. J Neurosci 31(1):164-171
Robustness of cortical sound encoding to synthetic and to real-world background noise

Alexander J. E. Kell (1,2), Erica Shook (1,2), Josh H. McDermott (1,2)

(1) Department of Brain & Cognitive Sciences, MIT
(2) Center for Brains, Minds, and Machines, MIT

In everyday listening, the sounds from sources we seek to understand are often embedded in background noise. This noise alters the pattern of spikes in the auditory nerve, often to a profound degree. In order to recognize sources of interest, the brain must to some extent become robust to the effects of these background noises. To study the neural basis of listening in real-world background noise, we measured fMRI responses in human auditory cortex to a broad set of thirty natural sounds, presented in quiet as well as embedded in thirty different everyday background noises (e.g., a bustling coffee shop, crickets chirping, heavy rain hitting pavement). We quantified the extent to which neural responses were robust to background noise by correlating each voxel’s response to the natural sounds in isolation with its response to those same natural sounds mixed with background noise. We normalized this correlation coefficient by the reliability of the voxel’s response measured across split halves of the data. This measure quantifies the extent to which a voxel’s pattern of response across natural sounds is the same when the natural sound is presented in quiet and when it is embedded in noise. Responses in anatomically-defined primary areas (TE 1.1 and 1.0) were substantially affected by background noise ($r^2 \sim 0.40$). However, voxel noise-robustness was substantially higher in non-primary areas, in some cases being hardly affected by the background noises ($r^2 \sim 0.85$). Mean responses in primary and non-primary regions were both only slightly lower in the presence of background noise, indicating that the difference in robustness between regions was not due to the background noises differentially suppressing responses in primary areas. Additionally, to enable comparisons with previous animal physiology work, we also examined the robustness of cortical responses to spectrally-shaped noise. The results illustrate a method for studying the noise robustness of neural responses throughout human auditory cortex.
Inactivation of primate dorsolateral prefrontal cortex during auditory and visual working memory

Shraddha Shah (1), Bethany Plakke (1), Theodore Lincoln (1), Katryna Kevelson (1), James Bigelow (2), Lizabeth M. Romanski (1)

(1) Department of Neuroscience, University of Rochester School of Medicine and Dentistry

(2) University of California, San Francisco

The prefrontal cortex (PFC) is an essential node in the process of working memory, as has been extensively established through neurophysiology, neuropsychology and neuroimaging studies, in both humans and nonhuman primates. A large proportion of the studies that have investigated the neural mechanisms of working memory, especially in nonhuman primates, have utilized only visual stimuli as the memoranda to infer neural processes. However, the frontal lobes are also important for language and communication, and therefore should demonstrate involvement in auditory working memory. Early studies demonstrated that large lesions which included portions of both dorsal and ventral PFC, impaired auditory discrimination in non-human primates. Furthermore, neurophysiological investigations have recorded neurons in DLPFC which were active during auditory working memory tasks. Lastly, studies and computational models of visual working memory, suggest that dorsolateral prefrontal cortex (DLPFC) may play a process-oriented role, irrespective of modality. Therefore, investigation of the prefrontal cortex in auditory working memory is important to establish a broader understanding of prefrontal working memory functions and whether they are ‘amodal’ in nature. Our previous research studies have established that neurons in the ventrolateral prefrontal cortex (VLPFC) are active during audiovisual working memory and that inactivation of VLPFC, with reversible cortical cooling, decreased performance of auditory working memory but not visual working memory. In the current study we asked if inactivation of DLPFC would show similar effects as our inactivation of VLPFC or if there would be task related effects unrelated to stimulus modality. We, therefore, inactivated the DLPFC in nonhuman primates while they performed auditory and visual delayed response tasks. Assessment of performance on each task during DLPFC inactivation will determine whether DLPFC plays a general role in working memory or a modality-specific function. These experiments will help us understand the differential (or possibly overlapping) contributions of dorsal and ventral prefrontal cortex to working memory across different modalities.
Neural substrates of individual differences in speech-in-noise understanding ability

Inyong Choi (1,2), Subong Kim (1), Adam Schwalje (2), Andrew Liu (2), Philip Gander (2), Youngmin Na (3), Jihwan Woo (3), Robert McMurray (1), Timothy Griffiths (4)

(1) University of Iowa
(2) University of Iowa Hospitals and Clinics
(3) University of Ulsan
(4) Newcastle University

Even young normal hearing listeners show great variance in their ability to understand speech with interfering background noise. The main aim of this study is to investigate the neural correlates of individual differences in speech-in-noise understanding. Cortical models of speech processing involve left-hemisphere inferior frontal gyrus (IFG). However, it is unclear how the IFG activity contributes to the behavioral variance in speech-in-noise performance. We tested primary auditory cortex and IFG activity in thirty normal hearing listeners and fifty cochlear implant users during a 4-AFC monosyllabic word-in-noise understanding task using cortical surface-constrained EEG source analysis. Our results showed better encoding of speech-acoustic features in high performers as reflected in stronger and earlier responses in the auditory cortex than those of low performers. High performance was also related to earlier left IFG activity, while low performance was linked with later and greater IFG activity in both hemispheres. This finding demonstrates that both auditory and bilateral prefrontal cortices are recruited during speech processing while differential timing and lateralization of IFG activity is correlated with behavioral performance. Prolonged IFG activity may also indicate exertion of more listening effort by low performers in noise.
Tracking the gradual formation and decay of auditory sensory memory using behavior and concurrent EEG recordings in macaque monkeys

Tobias Teichert (1)

(1) University of Pittsburgh

For several seconds auditory information is passively stored in auditory sensory memory. Despite the importance of auditory sensory memory for many aspects of auditory function, its neural mechanisms are still a matter of debate. However, it has been noted that the amplitude of the auditory evoked N1, which is reduced immediately after a tone has been processed, recovers back to baseline at the same rate at which information decays from auditory sensory memory. Here we tested the hypothesis that amplitudes of auditory evoked potentials (AEPs) elicited by a specific tone are smaller if that tone is encoded more strongly in auditory sensory memory.

To that aim we recorded AEPs from 32 cranial EEG electrodes while macaque monkeys performed a novel delayed pitch-discrimination task designed to track the dynamic formation and decay of auditory sensory memory. In the task, animals listened to sequences of standard tones and released a lever when they identified a pitch-deviant target tone. The stimulus-onset asynchrony (SOA) of consecutive tones varied randomly between 0.250 and 12 sec. The target could occur between sequence positions 2 and 13. The frequency-difference between standard and target (ΔF) varied between 0 and 1.2 octaves. On catch trials (ΔF=0) animals were rewarded for not releasing the lever.

Target detection rate increased with ΔF. The slope of the corresponding psychometric function was used to quantify discrimination performance as a function of SOA and the number of preceding standards. Preliminary data showed that discrimination performance gradually increased with repetition number, and decreased with SOA. Comparison to behavior in homolog signal detection tasks without sensory memory component suggests that changes of performance in the discrimination task reflect the gradual strengthening of sensory memory with repetition and its gradual decay during periods of silence.

The hypothesis that small AEPs are a marker of strong memory encoding thus predicted that AEP amplitude would be small for short SOAs and after many repetitions. Indeed, several AEPs such as the P31 and the N85 --the presumed monkey homolog of the N1-- were smallest for the shortest SOAs. However, contrary to the prediction AEP amplitudes generally increased with stimulus repetition. Both effects shared similar timing and topography with one key exception: between 40 and 60 ms after tone-onset fronto-central electrodes (human Fz homolog) encoded SOA, while the effect of repetition number was either completely absent or substantially weaker. Interestingly, ERPs at the same latency and topography were reduced by stimulus-specific adaptation in a passive listening task.
Taken together, these findings suggest a specific role for this fronto-central EEG component in stimulus-specific adaptation and sensory memory. However, additional quantitative analyses are needed to link this component more closely to performance in the delayed tone-discrimination task and single-cell responses in auditory cortex.
Cell-type-dependent modulation of auditory processing by locomotion in the inferior colliculus

Chenggang Chen (1), Sen Song (1)

(1) Tsinghua University

Cortical and thalamic sensory responses are known to be modulated by behavioral state. While the visual processing are enhanced by locomotion, the auditory processing are significantly suppressed. However, the effects of behavioral state on the responses of sensory midbrain is poorly understood.

Here we report on two-photon population recordings of auditory responses from cell-type-specific neurons in the dorsal auditory midbrain or inferior colliculus (IC) with head-fixed, free-running mouse.

Locomotion caused weak effects on the spontaneous activities of excitatory neurons. When stimulated with sound in locomotion, auditory spectral and temporal responses of more than half of excitatory neurons were significantly enhanced, along with increased response reliabilities and signal-to-noise ratio. Moreover, unlike the noise-correlation (tendency of neurons to respond together) that reduced by locomotion in the cortex, which is thought to facilitate sensory encoding, in the IC, the noise-correlation of excitatory neurons was increased by the locomotion state. Furthermore, this elevated correlated variability was found to increase the Fisher information (i.e., improve the population coding accuracy). In contrast, locomotion strongly activated most inhibitory neurons in the IC independent of sound stimulation, whereas the auditory activities were diversely suppressed by locomotion. At last, anatomical tracing revealed that excitatory neurons received more cortical feedback, neuromodulatory nuclei acetylcholine (ACh) and norepinephrine (NE) inputs than inhibitory neurons.

Therefore, subcortical IC auditory responses are also behavioral state dependent, but the modulation effects, population dynamics and neural circuit are different to the cortex and thalamus, a finding with implications for the neuroethological role that mediated by midbrain.
Homeostatic normalization of sensory gain in auditory corticofugal feedback neurons

Meenakshi M Asokan(1,2), Ross S Williamson(1,3), Kenneth E Hancock(1,2,3), and Daniel B Polley(1,2,3)

(1) Eaton-Peabody Laboratories, Massachusetts Eye and Ear Infirmary, Boston MA 02114

(2) Speech and Hearing Bioscience and Technology Program, Harvard Medical School, Boston MA 02115

(3) Department of Otolaryngology, Harvard Medical School, Boston MA 02114

Central auditory responses to suprathreshold sounds are paradoxically increased following a loss of cochlear afferent neurons. Central gain enhancement is more robust in the auditory cortex (ACtx) than the inferior colliculus (IC), suggesting a progressive increase in the expression of response enhancement along the central auditory neuroaxis. However, the ACtx also provides excitatory input back to the IC via a massive network of corticofugal projections. Little is known about the intrinsic, naturally occurring plasticity within the corticofugal feedback neurons themselves, due to technical challenges associated with targeted long-term recordings from distributed deep layer cortical neurons. Here, we describe a chronic widefield calcium axon imaging approach to track day-by-day dynamics in auditory corticocollicular (CCol) sound processing following auditory deprivation in adult mice. The hearing loss protocol used allowed us to image sound-evoked CCol responses at three phases: i) a stable baseline period, ii) a brief period of elevated thresholds immediately following noise exposure, and iii) a period with normal cochlear thresholds but a permanent loss of cochlear afferent synapses. We found that sound-evoked CCol responses were suppressed several hours after moderate intensity noise exposure but returned to baseline levels within 24 hours, a time point when cochlear thresholds were significantly elevated. CCol gain rose above baseline for several days following noise exposure before stabilizing at pre-exposure levels within 6 days. These findings demonstrate a homeostatic plasticity in the descending projections from the ACtx to the IC that can compensates both for a temporary elevation of cochlear thresholds and a permanent loss of cochlear afferent synapses. Future work will reveal how the plasticity of these auditory corticofugal projection neurons is coordinated with pre- and post-synaptic gain changes to shape adaptive and maladaptive changes in auditory perception following peripheral damage.
Optogenetic manipulation of inhibitory populations in auditory cortex during stimulus-specific adaptation

Tohar Sion Yarden (1), Adi Mizrahi (1), Israel Nelken (1)

(1) Department of Neurobiology, Silberman Institute of Life Sciences, and the Edmond & Lily Safra Center for Brain Sciences, Hebrew University of Jerusalem

Neurons in primary auditory cortex exhibit stimulus-specific adaptation (SSA), the decrease in responses to a common (standard) stimulus that does not generalize fully to another, rare stimulus (deviant). We investigate the role of inhibition in shaping SSA using optogenetic manipulation of the activity of specific inhibitory populations in auditory cortex. We used a viral vector for conditional expression of the Arch opsin in transgenic mice, and performed in vivo two-photon targeted loose-patch recordings, targeting both pyramidal neurons and opsin-expressing inhibitory neurons. We show that the optogenetic effects on the opsin-expressing neurons were sometimes non-trivial, e.g. increasing the sensory response rather than decreasing it, or decreasing it more strongly at lower light intensities. Optogenetic effects depended on the protocol of sensory stimulation and on the intensity of light stimulation, and were different for spontaneous activity and for different phases of the sensory-evoked activity. In addition, we relate the manipulation of inhibitory neuron activity to the responses recorded in pyramidal neurons, showing the contribution of inhibition to the excitatory cortical activity during SSA.
Cortical responses and functional connectivity derived from electrocorticography (ECoG) during speech in noise task

Andrew Liu (1), Phil Gander (2), Christopher Kovach (2), Hiroto Kawasaki (2), Matthew Howard (2), Tim Griffiths (3), Inyong Choi (4)

(1) Department of Otolaryngology-Head and Neck Surgery, University of Iowa, (2) Department of Neurosurgery, University of Iowa, (3) Newcastle University, (4) Department of Communication Sciences and Disorders, University of Iowa

Understanding speech in noisy environments can be quite challenging, especially so for hearing impaired listeners. Performance on speech in noise word recognition tasks cannot be predicted directly from an audiogram alone, and speech degradation has been recently shown to recruit inferior frontal gyrus (IFG) in addition to auditory cortical areas along the superior temporal gyrus (STG). Recent human non-invasive neuroimaging work suggests that STG activity represents stimulus-driven low-level perceptual processing while IFG is recruited for top-down feedback processing to resolve perceptual ambiguity resulting from stimulus degradation.

Using a speech in noise word recognition task based on the California Consonant Test (CCT), we collected intracranial electrocorticography (ECoG) data in five patients implanted with intracranial subdural and depth electrodes for clinical seizure monitoring. All subjects had normal hearing. During a 4AFC word recognition task subjects listened to CVC monosyllabic words presented in multitalker babble at +3 and -3 dB SNR. The answer choices were drawn from a set of matching rhyme or cohort words.

We examined the role of gamma and high gamma band responses in Heschl’s Gyrus (HG), STG, and IFG during the task. Additionally, we analyzed the functional connectivity within auditory cortex as well as between auditory cortex and IFG. These results reveal the role of IFG as well as the functional architecture of both auditory and frontal cortices during speech perception in difficult hearing situations."
Effects of propofol anesthesia on connectivity within the human auditory cortical hierarchy: An intracranial electrophysiology study

Matthew I. Banks (1), Kirill V. Nourski (2), Hiroto Kawasaki (2), Matthew A. Howard III (2)

(1) University of Wisconsin School of Medicine and Public Health
(2) The University of Iowa

While anesthetic action on subcortical sites likely contributes to loss of consciousness (LOC), there is an emerging consensus on the importance of disrupted cortical network connectivity, a model with extensive, but indirect, experimental support. Previously, we showed in murine core auditory cortex that feedback cortical pathways are differentially sensitive to suppression by general anesthetics compared to thalamocortical feedforward pathways. Whether this applies to the human cortical hierarchy is unclear. We addressed this issue using electrocorticographic (ECoG) recordings in neurosurgical patients during induction of propofol anesthesia.

Subjects were undergoing removal of intracranial electrodes placed to identify epileptic foci. ECoG recordings were made simultaneously with depth electrodes implanted in superior temporal plane and subdural grid electrodes implanted over the hemispheric convexity. We focused on analysis of resting-state data recorded in 5 minute blocks interleaved with responses to a local global deviant (LGD) paradigm. Data were collected during an awake baseline period and during induction of general anesthesia with incrementally titrated propofol infusion. Spectral analysis was performed using multitaper methods and was targeted to nodes in the auditory cortical hierarchy that were responsive during the LGD paradigm: posteromedial and anterolateral Heschl’s gyrus (PMHG, ALHG), planum temporale (PT), superior and middle temporal gyrus (STG, MTG), supramarginal gyrus (SMG), inferior and middle frontal gyrus (IFG, MFG). Connectivity analysis was targeted to these nodes, as well as applied to the entire recorded network. Functional connectivity (FC) was measured as weighted phase lag index (WPLI) in conventional frequency bands (delta to gamma).

In SMG, IFG and MFG, propofol anesthesia was associated with large, wideband increases in spectral power, with the largest increases in the alpha band, but modest or no changes in spectral power elsewhere. Under waking conditions, gamma FC was high locally within auditory cortex (PMHG, ALHG, PT), but was lower elsewhere. Alpha FC was lower in magnitude but more far ranging. Gamma FC was largely resistant to anesthesia, whereas alpha FC, especially between distant nodes, was enhanced under anesthesia.

Data demonstrate frequency band- and region-specific effects of general anesthesia on network connectivity along the human auditory cortical hierarchy. Modest effects on gamma-band WPLI are consistent with sparing of feedforward cortical connectivity.
Enhanced connectivity in the alpha band may contribute to the loss of consciousness during induction of general anesthesia.

This study was supported by grants NIH R01-DC04290, NIH R01-GM109086, NIH UL1-RR024979, NSF CRCNS-IIS-1515678 and the Hoover Fund.
Decomposition of the Neurophonic in Barn Owl Nucleus Laminaris

Paula T Kuokkanen (1,2,3), Anna Kraemer (3), Thomas McColgan (1), Catherine E Carr (3), Richard Kempter (1,2)

(1) Humboldt Universitaet zu Berlin
(2) Bernstein Center for Computational Neuroscience Berlin
(3) University of Maryland

In-vivo neural activity gives rise to trans-membrane currents that can be recorded as an extracellular field potential (EFP). These potentials are often challenging to interpret due to thousands of contributing sources. We aim at revealing the neural sources of the ‘neurophonic’. The neurophonic is a frequency following extracellular potential that can be recorded in the network formed by the nucleus magnocellularis (NM) and the nucleus laminaris (NL) in the brainstem of the barn owl. NL anatomy is well understood, and putative generators of the neurophonic are the activity of afferent axons from NM, the synaptic activation onto NL neurons, and spikes of NL neurons.

We recorded the neurophonic in response to binaural tones (3-7 kHz) around the recording site's best frequency, and varied the interaural time difference (ITD). The mean activity of the monaural inputs to NL does not change with ITD, although their relative phase does, causing cancellation or summation of input signals in the neurophonic, while the postsynaptic output of NL is strongly ITD tuned. Our extracellular recordings contained both of these signals, and we used pharmacological experiments as well as analytic and modelling techniques to decompose the neurophonic into its component sources.

Models predicted that the input axons from NM are sufficient as the main source of neurophonic, and comparison with experimental data showed that the inputs explained >95% of the strong high-frequency component (> 1 kHz) of the neurophonic and some of the low-frequency components (< 1 kHz). Pharmacological experiments with the AMPA-receptor antagonist NBQX revealed that synaptic EFPs in the barn owl NL are significant at frequencies < 500 Hz, but their overall contribution is quite small (< 1 % of the variance). The temporal dynamics of the evoked synaptic contributions to the EFP are consistent with synaptic short-term depression (STD).

Analyses of the broad-band power spectrum of the neurophonic allowed us to further separate the contributions of the presynaptic monaural inputs from the postsynaptic responses. We measured changes with ITD of the low-frequency component (LFc, 200-700 Hz) of the neurophonic. The spectrum of extracellularly recorded NL neurons' action potentials closely resembled this component. The contribution of the action potentials to this LFc ranged 0-50%, probably depending on how close the recording electrode was to the nearest NL neuron. Thus, the LFc reflects the contribution of action potentials initiated potentially from a single NL neuron.
We thus have decomposed the EFP in the barn owl nucleus laminaris for frequencies >200 Hz. Unexplained residuals of the neurophonic at < 100 Hz possibly originate from the ITD- and frequency-unspecific recurrent inhibition from the superior olivary nucleus or from the slow potassium currents on the somata of NL neurons, which serve to compensate for DC offsets in the synaptic current.
Dissociating task acquisition from expression during learning reveals latent knowledge

Kishore V. Kuchibhotla (1), Tom Anton Hindmarsh Sten (1), Eleni Papadoyannis (1), Srdjan Ostoji (2), Robert C. Froemke (1)

(1) NYU School of Medicine
(2) Ecole Normale Superieure

Performance on cognitive tasks is often used to measure intelligence yet remains controversial, largely because such testing is susceptible to contextual factors. To what extent does performance during learning depend on underlying knowledge or the testing context? Here, we report that acquisition and expression of task knowledge can be dissociated during learning in mice by manipulating the testing context. Strikingly, mice reveal a latent ability to perform a cognitively demanding task with individual differences in learning curves reflecting sensitivity to testing context rather than differences in underlying task knowledge. We trained mice on an auditory recognition task, and interleaved rewarded trials with unrewarded probe trials. Surprisingly, behavior on probe trials showed that animals understood the meaning of task stimuli at much earlier stages of task learning than suggested by their performance on reinforced trials. While learning curves on reinforced trials were highly variable across animals, probe learning curves were highly stereotyped. To mechanistically account for these observations, we constructed a neural circuit model in which acquisition and expression were explicitly dissociated, with task acquisition based on reward-driven plasticity of sensorimotor projections (representing task knowledge) and expression modulated by context in a decision circuit. Fitting the model to experimental data demonstrated that modulation of feedforward inhibition accounted for performance differences between contexts as well as variability between animals. Individual performance variance therefore emerges more from contextual factors than from differences in underlying rates of associative learning. Overall, testing appears to impose critical limitations on the expression of underlying task knowledge and may confound interpretation of performance.
Synergistic shifts in AMPA and GABAA receptor transcripts underlying recovered auditory processing in the adult auditory cortex following damage to the inner ear.

Pooja Balaram (1,2), Daniel B. Polley (1,2)

(1) Eaton-Peabody Laboratories, Massachusetts Eye and Ear Infirmary
(2) Department of Otolaryngology, Harvard Medical School

Damaging hair cells in the inner ear depresses sound-evoked activity at early stages of the auditory pathway but causes a paradoxical increase in activity at the level of auditory cortex (ACtx). Increased ‘central gain’ is associated with variable levels of recovered cortical processing; some mice regain normal auditory thresholds while other mice show no recovery of auditory processing, despite an equivalent loss of cochlear afferent synapses and auditory brainstem responses.

How do auditory cortical neurons recover responses to sound following peripheral injury? We hypothesized that homeostatic regulation of AMPA and GABAA postsynaptic receptors would contribute to central gain after cochlear insult, and might explain functional outcomes for sensory processing between individuals. Further, homeostatic regulation of postsynaptic receptors could differ between excitatory and inhibitory cortical neurons, and the opposing directions of receptor shifts between these two populations could work synergistically to maximally increase central gain.

To address these questions, we measured sound-evoked responses in awake, adult head-fixed mice using widefield and two-photon imaging of cortical neurons expressing GCaMP6s under control of the CAMKIIa promoter. We unilaterally ablated cochlear hair cells via cochlear perfusion with distilled water, tracked daily changes in sound-evoked responses across ACtx for five days, and documented variable levels of damage and functional recovery between mice. We then quantified transcriptional levels in genes that encode AMPA and GABAA receptor subunits in individual auditory cortical neurons. Changes in the transcription of these subunits reflect subsequent changes in the translation of AMPA and GABAA receptors, thus providing a correlative measure of synaptic scaling in ACtx.

Cortical neurons were chemotyped by documenting the expression of VGLUT1 or VGLUT2 mRNA (excitatory neurons), or VGAT mRNA (inhibitory neurons). Expression levels of Gria2 mRNA, which encodes the GluA2 subunit of AMPA receptors, and Gabra1 mRNA, which encodes the alpha1 subunit of GABAA receptors, were separately quantified in VGLUT-positive and VGAT-positive neurons across ACtx layers (L)2/3, L4 and L5/6. Following cochlear injury, Gria2 and Gabra1 transcripts were significantly shifted relative to sham-operated controls, with complementary transcriptional shifts between VGLUT-positive and VGAT-positive neurons. These synergistic transcriptional shifts across local cortical networks may provide the means to
increase the gain on reduced afferent inputs and recover normal activity levels in the ACtx.
Cell type-specific long-range connections of the higher-order auditory thalamus

Dongqin Cai1, Yin Yue1, Xin Su2, Yiwei Wang1, Miaomiao Liu1, Ling You1, Fenghua Xie1, Fei Deng2, Feng Chen2, Minmin Luo3, Kexin Yuan4*

(1) Department of Biomedical Engineering, School of Medicine, IDG/McGovern Institute for Brain Research, Tsinghua University, Beijing, 100084, China.

(2) Beijing KLSBDPA Key Laboratory, Department of Automation and CBICR Center, Tsinghua University, Beijing, 100084, China.

(3) National Institute of Biological Sciences, Zhongguancun Life Science Park 7 Science Park Road, Beijing 102206, China.

(4) Department of Biomedical Engineering, School of Medicine, IDG/McGovern Institute for Brain Research, Center for Brain-Inspired Computing Research, Tsinghua University, Beijing, 100084, China.

*Correspondence to: Kexin Yuan, Phone: 86-10-62783759, Fax: 86-10-62773380, Email: kexinyuan@tsinghua.edu.cn

These authors contributed equally to this work.

Although many traditional anatomical studies have shown the connectivity between the higher-order auditory thalamus and other brain areas, data showing cell type-specific whole-brain connectivity for different subdivisions has been lacking. Previous immuno-histological studies demonstrated localized expression of two types of calcium binding proteins, calbindin (CB) and calretinin (CR), in the higher-order auditory thalamus. Using anterograde and retrograde viral tracing in CB- and CR-Cre mice, we found largely distinct cortical and subcortical connection patterns for different subdivisions of higher-order auditory thalamus in both CB- and CR-cre mice. PIN/PP formed widely reciprocal connection with brain areas concerning multisensory information processing, certain behavior controlling and modulatory neurotransmitter releasing. In contrast, MGBd and MGBm connected with fewer regions and most of them were involved in auditory or visual information processing. Further examination of connection pattern within some nuclei also showed marked difference between PIN/PP and MGBd, such as SC, AC and CPU. Meanwhile, no significant difference in connection pattern across all subdivisions was found between CB- and CR-cre mice groups. Our results show that, due to the localized expression of CB and CR in the higher-order auditory thalamus, diffusion of traditional tracers into the first-order part is not an issue anymore. Most importantly, our results suggest that CB- and CR-Cre mice can be used as a powerful tool to explore the functions of specific pathways between the higher-order auditory thalamus and other brain areas. These anatomical findings and potential functional studies will deepen our understanding of the role of the high-order auditory thalamus in information processing and animal behavior.
Eye movement-related eardrum oscillations (EMREOs): a biomarker for visual-auditory spatial integration in the auditory periphery?

David L. K. Murphy (1), Kurtis G. Gruters (1), David W. Smith (2), Christopher A. Shera (3), Jennifer M. Groh (1)

(1) Duke University
(2) University of Florida
(3) University of Southern California

Visual and auditory signals that arise from coincident locations in space often fuse to produce a combined visual-auditory percept, such as lip reading cues that facilitate speech comprehension. A critical problem for this process is that the visual and auditory systems employ different methods of determining the locations of stimuli. The visual system uses retinal activation location and the auditory system evaluates interaural timing and intensity differences as well as spectral cues. In species that move their eyes with respect to the head, there is no fixed relationship between these cues, and representations must be adjusted with each change in eye position if the brain is to determine whether a visual and an auditory cues spatially coincide.

Eye movements have been shown to modulate auditory activity in several brain regions such as the inferior colliculus, auditory cortex, parietal cortex, frontal eye fields, and the superior colliculus, but where eye movement-related signals first contribute to auditory processing is not known. One possibility is the auditory periphery. Outer hair cells, or motile neurons in the cochlea, and the middle ear muscles are under descending control from the brain, providing a possible route of transmission of information about eye movements to the auditory periphery.

To test this hypothesis, we used a microphone to record movements of the eardrum in human subjects (n=10, 16 ears) performing saccades to visual targets. We found that the eardrums oscillated in conjunction with eye-movements, despite the absence of any delivered sound. These oscillations began at least 10ms before the start of any given saccade. The initial phase and magnitude of these eye-movement related eardrum oscillations (EMREO) was dependent on saccade origin, direction, and length. This relationship suggests it contains information necessary to facilitate visual-auditory spatial integration reaches the auditory periphery. EMREOs are a candidate biomarker that may prove useful for assessing the cause of multisensory integration deficits in affected populations, although the exact means by which it does so remains uncertain.
Distinct sensory and extra-sensory processing differences in two types of deep layer auditory cortex projection neuron

Ross S Williamson (1,2), Daniel B Polley (1,2)

(1) Eaton-Peabody Laboratories, Massachusetts Eye and Ear Infirmary
(2) Department of Otolaryngology, Harvard Medical School

Neurons in layers (L) 5 and 6 of the auditory cortex (ACtx) give rise to a massive subcortical projection that innervates all levels of the central auditory pathway as well as non-auditory areas including the amygdala and striatum. L5 and L6 neurons feature distinct morphology, connection patterns, intrinsic membrane and synaptic properties, yet little is known about how these differences relate to sensory selectivity in vivo. Here, we performed cell-type-specific single-unit recordings and wide-field calcium imaging from two classes of ACtx L5 and L6 projection neurons; L5 corticocollicular neurons (L5CC), and L6 corticothalamic neurons (L6CT).

We first used an antidromic optogenetic ‘phototagging’ approach to isolate individual L5CC and L6CT units on high-density multichannel probes in awake, head-fixed mice. We constructed a bank of complex stimuli by endowing noise tokens with randomly chosen acoustic parameters, and then quantified each neuron’s lifetime sparseness. We found that L5CC neurons had a lower lifetime sparseness index, indicative of reduced stimulus selectivity and a broader response distribution than L6CT neurons. Linear spectrotemporal receptive field fits were also able to explain a higher percentage of response variance in L6CT neurons, indicating a higher degree of linearity in their responses when compared to L5CC neurons. Recent studies have shown that an animal’s ‘internal state’, as indexed through locomotion or pupil diameter, exerts a powerful influence over ACtx excitability. We expressed GCaMP6s in L5CC and L6CT neurons and noted that bouts of locomotion were associated with opposite effects on the wide-field calcium signal: a decrease in L5CC neurons but a surprising increase in L6CT neurons. Collectively, these studies show that each class of deep-layer projection neuron performs distinct operations on internal and external signals, which likely impart distinct effects on their subcortical targets.
The brain’s ability to selectively isolate and enhance a single speech stream from multiple competing sources allows us to successfully communicate in everyday life. In recent years, a great deal of progress has been made towards understanding this complex process. This has, for the most part, been provided through the use of ‘decoders’ which transform the recorded neural activity into an estimate of the heard acoustic envelope (Mesgarani and Chang 2012, Ding and Simon 2012). Crucially, these methods allow us to examine the neural dynamics of attention using natural, continuous speech. However, research on this topic to date has been constrained somewhat by the fact that it focuses purely on auditory speech. And we know from our everyday experience that listeners will often look at the talker’s face in an effort to aid comprehension -- especially in a noisy environment. Indeed, recent work on audiovisual speech has quantified this benefit and shown it is present for speech in quiet (Crosse et al., 2015), and even more so in the presence of background noise (Crosse et al., 2016). Importantly however, listeners will also often switch their attention to a different audio source (eavesdropping), while continuing to look at the face of their conversational partner. Understanding how attention operates under both of these conditions is central to advancing our knowledge on the interactions between top-down attentional selection and the bottom-up visual influence on speech processing. Here, we present subjects with congruent audiovisual (AV) speech in the presence of a competing audio (A) stream. Subjects are required to attend to one speech stream (condition 1: attend A; condition 2: attend AV) while ignoring the other. Our results show that combining indices of both audio and visual speech processing provides a robust attention decoder which can successfully operate under both of these conditions. In addition, we find that contributions from occipital regions aids decoding of attention to AV speech, whereas these regions are heavily down weighted when decoding attention to A speech. We also examine the underlying temporal dynamics of the interaction between visual speech and audio speech processing, and the extent to which this is influenced by attention. This work may lead to important neuroscientific insights on the interaction between attention and multisensory integration and aid progress towards the goal of an attentionally steered hearing aid.
Auditory cortex is essential for mammals, including rodents, to behaviorally detect temporal cues in sound but it remains unclear what different cortical fields contribute to temporal cue discrimination (Threlkeld et al., 2008; Lomber & Malhotra, 2008). Previously, we found temporally precise spiking patterns change proportionally with temporal shape cues in the sound envelope for spike outputs from primary (A1) and non-primary Ventral and Suprarhinal (VAF, SRAF) auditory fields of the rat (Lee et al., 2016). Here we examine how these temporal spiking patterns may be used to discriminate and classify sound shape cues. Confirming and extending prior studies in midbrain and A1 (Phillips et al. 2002; Heil 2003), we find the temporally precise output provides a robust signal that changes with sound shape in non-primary cortices. Moreover, we find temporally precise spiking and not spike rate output provides the best signal for discrimination and classification of sound in cortex as described previously for midbrain and A1 (Schneider and Woolley 2010; Malone et al. 2015). A compelling new contribution from this study is the finding that time scales for temporally precise output and optimal discrimination increase in rank order with: A1 < VAF < cSRAF. In addition, we find that response peak delay and response duration mirror changes in sound shape parameters and related temporal spiking patterns provide an effective neural code for sound shape in all three cortical fields. When onset spikes are removed we find persistent spiking responses in non-primary cortex provide a temporal signal that can discriminate and classify sound shape. This is not the case for A1 neurons. The dependence on sound shape is quite different for response peak delay versus duration, suggesting these temporal response patterns may provide unique neural codes in A1 and non-primary cortex. Together these results support the concept that these three auditory cortical fields provide temporally precise output on complementary time scales that could underlie the ability of animals to discriminate and classify the temporal shape of the sound envelope.
Representations of amplitude modulations in auditory onsets, ramp tones, and speech in the human superior temporal gyrus

Yulia Oganian (1), Edward F. Chang (2)

(1) University of California, San Francisco
(2) University of California, San Francisco

Making sense of complex auditory inputs requires temporally precise parsing of single events out of the auditory stream. Auditory event onsets are typically marked by an increase in amplitude, and previous animal and human studies identified the dynamics of amplitude rise at onset as a central feature encoded throughout the subcortical auditory pathway. Amplitude modulations, however, also mark changes within ongoing sounds, e.g. in the speech envelope. Yet, it is unknown how the human auditory cortex differentiates between amplitude rises in silence and within an ongoing sound, and whether onset encoding can account for tracking of the speech amplitude envelope.

To address this, we designed tone stimuli containing amplitude ramps, rising from silence (RfS condition), or from an amplitude baseline (RfB condition). The sound intensity of ramps increased linearly to peak amplitude and returned to silence/baseline, with varying rates of amplitude change. We recorded local field potentials using intracranial multi-electrode arrays placed over the temporal lobes of five patients undergoing evaluation for epilepsy neurosurgery, as they passively listened to the tones. In both conditions, ramps elicited transient responses in the high gamma frequency range (HG, 70-150 Hz) in posterior (p) and middle (m) superior temporal gyrus (STG), with larger HG amplitudes for fast-rising ramps. We observed a striking double dissociation of response types: pSTG encoded the rate of amplitude change in the RfS condition only, whereas mSTG encoded the rate of amplitude change in the RfB condition only. Crucially, the rate of amplitude envelope modulation in continuous speech was also represented in mSTG, but not in pSTG.

Our results reveal functionally and spatially distinct representations of sound onsets in silence and in background along STG. Moreover, our data suggest that speech amplitude tracking in mSTG may rely on the same neural mechanisms as the encoding of onsets in background.
Cortical representations of attention to pitch or timbre

Emily J. Allen (1), Philip C. Burton (1), Cheryl A. Olman (1), Andrew J. Oxenham (1)

(1) University of Minnesota

Pitch and timbre are two primary dimensions of auditory perception that are generally considered to be independent. However, under some circumstances they have been shown to interact with one another. For example, an increase in the fundamental frequency of a sound, the nearest physical correlate of pitch, can be confused for an increase in spectral centroid, a physical correlate of brightness (i.e., an attribute of timbre), and vice versa. An earlier passive fMRI study indicated that pitch and brightness are processed in largely overlapping regions of the auditory cortex, but their patterns of activation are separable via multi-voxel pattern analysis (MVPA). The aim of the current fMRI study was to determine whether allocating attention to one dimension or the other would make their cortical representations more spatially distinct. We collected BOLD data at 3T while participants listened to pairs of tones varying in either pitch, timbre, or both, and judged which tone had higher pitch or brighter timbre. While task performance was high in all conditions, subjects were significantly worse at performing the task when both dimensions were varying, confirming previous behavioral findings. Consistent with our previous results, we found that pitch and timbre engaged a common set of auditory regions with no clear distinctions at a univariate level. Furthermore, we found that increased attentional demands in the conditions in which both sound dimensions varied led to increased engagement of frontal regions, including the inferior and middle frontal gyri, compared to conditions in which just one dimension varied. However, heightened attentional demands did not appear to improve the separability of the neural representations of pitch and timbre at either univariate, or multivariate levels. These results suggest that the computations underlying pitch and timbre perception are subserved by overlapping populations, likely on a scale that is difficult to disentangle without higher resolution and/or more sophisticated computational encoding models.
Behavioral significance and auditory cortical representation of harmonic vocalizations

Nina LT So (1,2)
Sarah MN Woolley (1,3)

(1) Zuckerman Institute, Columbia University
(2) Doctoral Program in Neurobiology and Behavior, Columbia University
(3) Department of Psychology, Columbia University

Vocal communication relies on the brain’s ability to detect and process vocal sounds in the environment. Auditory systems that are optimized for vocalization processing must robustly encode acoustic features that distinguish vocalizations from other sound categories. One acoustic hallmark of speech and other animal vocalizations is harmonic structure. Sounds that are harmonic contain energy concentrated at integer multiples of a fundamental frequency, and higher harmonicity correlates with greater pitch salience. Our studies address two main questions. First, how does variation on harmonic structure alter the behavioral significance of vocal sounds? Second, where does selectivity for harmonic structure arise in the auditory processing pathway? We use zebra finches (Taeniopygia guttata) to address these questions. These songbirds use harmonic calls and songs for social communication and have well-defined, experimentally accessible auditory systems.

Zebra finches exchange distance calls, a highly harmonic class of vocalizations, to maintain contact with each other when visually separated. Using a behavioral paradigm in which we present distance calls and assess birds’ vocal responses to these stimuli, we found that behavioral responsiveness is sensitive to the degree of harmonic structure in stimulus calls; birds respond less to spectrally-degraded versions of calls compared to natural versions of the same calls. To test whether the responses of auditory cortex (AC) neurons are similarly sensitive to harmonic structure, we recorded the responses of single AC neurons to natural and spectrally-degraded distance calls, in awake, restrained zebra finches while they listen to the calls. We recorded single unit responses from: 1) the intermediate region (Field L2), which receives information from the thalamus; 2) the superficial regions (Field L1/CM), which receive input from the intermediate region; and 3) the deep region (Field L3), which receives input from both intermediate and superficial regions. Preliminary results suggest that neuronal responses in the intermediate and superficial regions do not differ with variations in calls’ harmonic structure. In contrast, deep region neurons respond more to calls with high harmonic structure than to spectrally-degraded versions of those calls. Together, results suggest that harmonic structure contributes to the behavioral salience of vocal communication sounds, and that neural selectivity for this feature arises in the deep region of primary AC.
Anatomical Evidence for Parallel Processing of Auditory Information in the Songbird Brain

Mark J Basista (1,2), Sarah Woolley (1,2)

(1) Zuckerman Institute, Columbia University
(2) Department of Psychology, Columbia University

Vocal communication relies on the auditory system’s accuracy in transforming neural responses to acoustic features into representations of social information. While the evolutionary conservation of auditory pathways in the vertebrate brainstem is well understood, recent studies show parallels between avian and mammalian auditory pathways in cortex. Like mammalian cortex, the major processing pathway in avian auditory cortex (AC) is from the thalamo-recipient, intermediate region (Field L2) to the superficial regions (Field L1 and caudal mesopallium), and from superficial regions to the deep output region (Field L3). In the songbird AC, other connections among these regions have been shown. For example, Field L1 is reciprocally connected to the caudal mesopallium (CM) and Field L3 receives input from CM and the auditory thalamus. While the major pathways among songbird auditory cortical regions are well established, little is known about the specific cell-to-cell connections between regions. The interconnectivity of AC regions in songbirds suggests parallel processing of auditory information. If there are parallel processing streams in songbird AC, then discrete neural populations within a region should project to discrete neural populations in downstream regions. High resolution mapping of neural projections in AC could provide insight into the mechanisms whereby complex tuning properties arise in neurons that receive input from neurons with simple tuning properties, a feature common to birds and mammals. Here, we used tract-tracing fluorescent dyes to test the hypothesis that projections between auditory regions are organized into multiple, distinct pathways. We used zebra finches (Taeniopygia gutatta) because they rely on auditory processing of complex vocalizations (calls and songs) for social behavior and because the response properties of neurons in each region of AC are well described. In the zebra finch AC, single neuron responses to sounds are progressively more complex and nonlinear between the intermediate and deep output regions. Selectivity for song over other sounds emerges along this pathway. We made micro-injections of dyes in separate locations of the intermediate and deep output regions of AC, and located specific projection neurons and their spatial mapping in upstream regions. Preliminary results suggest that largely distinct neural populations project from thalamus to cortex and from one cortical region to another depending on spatial location within a region. Multiple separate connections between regions could provide a framework for parallel processing in the neural circuitry that underlies vocal communication.
43 Nonlinear population cortical responses after midbrain stimulation in the auditory colliculo-thalamocortical mouse brain slice.

Baher A. Ibrahim (1,2), Daniel A. Llano (1,2)

(1) The department of Molecular and Integrative Physiology, University of Illinois at Urbana-Champaign, Illinois, USA

(2) Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Illinois, USA

The auditory colliculo-thalamocortical mouse brain slice, developed by our laboratory, retains synaptic connectivity between inferior colliculus (IC), medial geniculate body (MGB), and auditory cortex (AC). Such connections in a brain slice preparation enable the detailed examination of the role of the MGB in transmitting signals from the IC to the AC. Using flavoprotein autofluorescence (FA) to measure the connectivity between circuit components, we recently observed all-or-none cortical FA responses without any change in IC and MGB responses following IC stimulation. The frequency of the missing FA cortical responses increased with decreasing current amplitude and inter-stimulus-interval. In contrast, direct stimulation of MGB or the subcortical white matter produced only linear cortical responses. Bath perfusion of gabazine, GABAA receptor blocker, was capable of retrieving the missing cortical FA responses. The preliminary data showed that focal injection of gabazine into MGB showed the same effect of the drug to retrieve the missing FA cortical responses and linearize cortical and MGB responses, but it failed to show the same effect when injected into auditory cortex. Current clamp whole cell recording from layer 4 or 2/3 showed similar all-or-none responses to what was described using FA, and voltage-clamp recording did not reveal a source of inhibition responsible for the missing responses. Current and voltage clamp whole cell as well as cell attached paired recording of MGB cells showed that not all MGB cells fire or receive inhibitory inputs following IC stimulation which could indicate population coding within MGB neuronal ensembles. These data suggest that the thalamus may recruit cortical ensembles rather than linearly encoding ascending stimuli, and that GABAergic inhibition may play a role in selecting cortical ensembles for activation.
Reticulothalamic and intrareticular synaptic microarchitectures determine oscillatory and propagative properties of thalamocortical waves

Jeffrey W. Brown (1), Aynaz Taheri (2), Robert V. Kenyon (2), Tanya Berger-Wolf (2), Daniel A. Llano (1)

(1) University of Illinois at Urbana-Champaign
(2) University of Illinois at Chicago

Recent experimental observations and modeling studies support both the existence and potential functional significance of ‘open-loop’ thalamo-reticulo-thalamic synapses, in which neurons from the thalamic reticular nucleus (TRN) are not reciprocally excited by the thalamocortical (TC) neurons they inhibit. We previously demonstrated that simulated open-loop thalamocortical networks are able sustain the propagation of spindle-like waves across the length of the network at a constant velocity. In the present study, we examine how systematically varying 14 synaptic weights within a simulated 9-neuron thalamocortical network alters the oscillatory and propagative properties of the waves the network supports. Utilizing regression algorithms, we demonstrate that specific individual synapses and synaptic clusters, comprising variously open- and closed-loop thalamo-reticulo-thalamic and intrareticular (GABAergic and electrical) microarchitectures within the network, are strongly correlated, both linearly and nonlinearly, with oscillatory and propagative wave dynamics. These findings are related to the thalamocortical physiology underlying both normal and pathological processes and compared to exclusively closed-loop models capable of approximating the same phenomena.
Dissociation of stimulus and outcome expectations in perceptual decision making

Akihiro Funamizu (1), Fred Marbach (1,2), Anthony M Zador (1)

(1) Cold Spring Harbor Laboratory

(2) Watson School of Biological Sciences

In perceptual decision making, humans and animals use the prior knowledge of stimulus probability and reward outcome to optimize their behavior. Although a series of studies have shown the neural substrate of outcome expectation in cortico-basal ganglia circuit, the neural substrate of stimulus expectation is unclear. To dissociate the two neural circuits, we are training head-fixed mice on an auditory frequency discrimination task (based on Marbach and Zador, bioRxiv 2016), in which either (i) the stimulus probability or (ii) the reward size for category A and category B trials changes in blocks, and analyze the choice behavior with a belief-updating reinforcement learning model.

We found that both the stimulus probability and reward amount biased mice toward choices associated with high-probability stimuli or large reward, respectively. In the stimulus probability task, a belief updating model, which recursively updated the probability of stimulus category, predicted the biased behavior better than a reinforcement learning model, which updated the expected outcome of each category. In contrast, in the reward amount task, the reinforcement learning model predicted the behavior better than the belief updating model. This suggests that mice updated their stimulus expectation in the stimulus probability task, while they updated their outcome expectation in the reward amount task. We expect that our paradigm combined with two-photon calcium imaging will help dissociate the neural circuits for stimulus and outcome expectations in the auditory cortex.
Modularity of intrinsic inputs within the lateral cortex of the mouse inferior colliculus

Alexandria M.H. Lesicko (1 and 2) and Daniel A. Lllano (1, 2 and 3)

(1) Neuroscience Program, University of Illinois at Urbana-Champaign
(2) Beckman Institute, University of Illinois at Urbana-Champaign
(3) Department of Molecular and Integrative Physiology, University of Illinois at Urbana-Champaign

The lateral cortex of the inferior colliculus contains a network of modules characterized by dense labeling for glutamic acid decarboxylase-67 (GAD-67) and other neurochemical markers. Previous studies from our laboratory have shown that the extrinsic sensory inputs to the lateral cortex are patterned: somatosensory inputs terminate within these neurochemical modules, while auditory inputs target the extramodular regions of the lateral cortex. While the topography of extrinsic inputs to the lateral cortex is well defined, the degree and directionality of intrinsic connectivity between neurons in this region remains unknown. In the present study, we sought to characterize the intrinsic inputs to neurons in the lateral cortex of the mouse inferior colliculus and determine if these connections also exhibit modularity. Experiments were performed in brain slices from the GAD-67-GFP knock-in mouse, in which modular and extramodular areas of the lateral cortex can be clearly distinguished. GAD-67-positive and GAD-67-negative cells in both regions were recorded from in either a single or dual-channel whole-cell voltage clamp configuration while potential pre-synaptic sites throughout the ipsilateral colliculus were stimulated using laser photostimulation of caged glutamate. Preliminary results include a heterogeneous set of input maps. Pre-synaptic stimulation generated both excitatory and inhibitory responses in cells within the lateral cortex; excitation resulted predominately from direct stimulation of the recorded cell, while strong synaptic inhibition was seen for the majority of cells. Neurons within modular regions of the lateral cortex received inputs primarily from presynaptic sites within a module, while extramodular cells were predominately activated from sites outside of the modules. These preliminary data indicate that the intrinsic connections within the lateral cortex may also exhibit modularity that is predicted by the underlying neurochemical modularity within this structure. Potential interconnectivity between modular and extramodular regions of the lateral cortex could have important implications for processing of multisensory information in the lateral cortex.
Tuning of cortical gain, sound frequency selectivity and auditory-driven behaviors by synaptically released zinc

Manoj Kumar*(1), Charles T. Anderson*(1), Shanshan Xiong(1), and Thanos Tzounopoulos(1)

(1)Department of Otolaryngology, University of Pittsburgh, Pittsburgh, PA 15261

* These authors contributed equally to this work.

In many excitatory synapses, mobile zinc is found within glutamatergic vesicles and is coreleased with glutamate. Ex vivo studies established that synaptically released (synaptic) zinc inhibits excitatory neurotransmission at lower frequencies of synaptic activity but enhances steady state synaptic responses during higher frequencies of activity. However, it remains unknown how synaptic zinc affects neuronal processing in vivo. Here, we imaged the sound-evoked neuronal activity of the primary auditory cortex in awake mice, and discovered that synaptic zinc enhanced the gain of sound-evoked responses in CaMKII-expressing principal neurons, but it reduced the gain of parvalbumin- and somatostatin-expressing interneurons. This modulation was sound intensity-dependent and, in part, NMDA receptor-independent. Moreover, synaptic zinc sharpened the sound frequency selectivity of CaMKII-expressing principal neurons, but it broadened the sound frequency selectivity of somatostatin-expressing interneurons. In the absence of cortical synaptic zinc, mice showed reduced ability in detecting behaviorally meaningful signals from their auditory environment, such as changes in sound frequencies. By establishing a previously unknown link between synaptic zinc and gain control of auditory cortical processing, our findings advance understanding about cortical synaptic mechanisms and create a new framework for approaching and interpreting the role of the auditory cortex in sound processing.

Supported by NIDCD, R-01 DC007905 award to TT
Optimal features for auditory recognition

Shi Tong Liu (1), Xiaoqin Wang (2), Srivatsun Sadagopan (1)

(1) University of Pittsburgh
(2) Johns Hopkins University

A central challenge in auditory neuroscience is to understand how observed patterns of neural activity in the auditory system relate to behavior. For example, neurons in primary (A1) as well as higher auditory cortical areas exhibit highly nonlinear and surprisingly specific tuning properties, but our understanding of these responses is only at a descriptive level, and the critical question of how these responses might support behavior remains unresolved. Here, we show that nonlinear A1 responses encode essential features for the classification of ethologically-relevant sounds such as conspecific vocalizations (calls). In vocal animals, increasing neural resources are committed for the processing of calls as one ascends the auditory processing hierarchy. Therefore, the categorization of call types is a reasonable computational goal for the auditory cortex in these animals. We asked, using a theoretical information maximization approach, how this goal can be best accomplished. We used marmoset vocalizations as our experimental model. First, we transformed the vocalizations into spectrotemporal patterns of auditory nerve activity (cochleagrams) using a highly realistic model of the auditory nerve. Based on an earlier model for visual classification, we then randomly generated a large number of features, or spectrotemporal snippets, from these cochleagrams. We used a greedy-search algorithm to choose the most informative and least redundant feature set for call categorization. We found that call categorization could be accomplished with high accuracy using just a small number of features. Highly informative features tended to be of intermediate size and complexity. Most interestingly, the responses of model feature-selective neurons predicted nonlinear neural responses in marmoset A1 in astonishing detail. These results demonstrate that the auditory cortex uses a mid-level feature based strategy for the recognition of complex sounds. These results further suggest that the tuning properties of neurons in higher auditory cortical stages are likely the result of goal-directed optimization.
Humans and vocal animals use vocalizations to communicate and interact with members of their species. Real-world environments add noises, echoes and other sounds to the intended message, degrading its acoustic content. However, we can maintain stable sound perception independent of listening conditions. We aim to determine the neural mechanisms by which stable sound perception can be achieved.

To address this question in the context of natural behaviors, we use Guinea pig (GP) vocalizations as an experimental model. Previous studies in GPs have shown that at the level of the inferior colliculus and thalamus, few neurons show selective responses for individual vocalization categories. In primary and secondary cortical areas, more neurons become selective for particular vocalization categories. It is not known, however, at which stage of the auditory hierarchy this selectivity arises, and how it is preserved or changed in the presence of real-world distortions. Here, we first tested if GPs can perceive vocalizations presented in a wide range of noisy environments using pupillometry as a behavioral readout. This allowed us to determine the GP’s threshold for detecting a vocalization in noise. We then recorded single-unit activity in the medial geniculate body (MGB) and auditory cortex (A1) of awake GPs passively listening to vocalizations in different listening conditions. We discovered that neurons in MGB and thalamorecipient A1 layers (A1 L4) have low selectivity for vocalization categories and are more susceptible to acoustic distortions. In contrast, superficial layers of A1 (A1 L2/3) were highly selective for vocalizations and more invariant to distortion. These data demonstrate that both vocalization selectivity and invariance to listening conditions co-emerge in A1 L2/3. These results suggest that a dense representation of complex sounds in A1 L4 is transformed into an invariant and sparse representation in A1 L2/3.

Grant/Other Support: Pennsylvania Lions Hearing Research Foundation, Samuel and Emma Winters Foundation.
Rapid cortical adaptation to changing noisy conditions

Bahar Khalighinejad (1,2), Jose L. Herrero (3,4), Ashesh D. Mehta (3,4), Nima Mesgarani (1,2)

(1) Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY
(2) Department of Electrical Engineering, Columbia University, New York, NY
(3) Hofstra Northwell School of Medicine, New York City, NY, United States
(4) The Feinstein Institute for Medical Research, New York City, NY, United States

Humans are experts at perceiving speech, even in the presence of interfering sound sources. Recent studies suggest that this perceptual ability may arise from a noise-invariant representation of speech in the auditory brain, resulting in several computational models to explain the robust encoding of stimulus. However, little is known about the dynamic properties of neural networks as the auditory system adapts to new noisy conditions.

To understand the adaptation process in human auditory cortex during varying additive distortions, we recorded intracranial electroencephalography signals from six human subjects who were implanted with high-density electrodes as a part of their clinical procedure. Subjects listened to continuous speech corrupted by additive background noise which changed randomly every three or six seconds to one of the four conditions of clean, bar, jet, and city noise, each of which have unique spectro-temporal properties. Analysis of the neural responses to the speech signal revealed a transient adaptation interval which lasts between 200 to 800 ms after the onset of noise change. The population coding of the stimulus revealed that spectral properties of the noises emerge during the adaptation, and they subsequently fade resulting in a noise-invariant post-adaptation response which has the properties of the clean speech. Moreover, we show that different regions of auditory cortex have unique adaptation patterns. Specifically, electrodes at higher auditory areas such as superior temporal gyrus show a stronger adaptive response in transition to the clean condition, while primary auditory areas such as medial Heschl’s gyrus are more adaptive to noisy conditions. Within primary areas, medial versus lateral regions show different adaptation patterns upon their tuning properties.
Early effects on subcortical processing of sound with learned extinction of auditory associative memory in a rodent model

Elena K. Rotondo (1), Alyssa Rodriguez (2), Kasia M. Bieszczad (3)

(1) Rutgers University
(2) Rutgers University
(3) Rutgers University

Representational plasticity in the auditory system is frequently studied in the context of long-term auditory memory and, in the auditory cortex, is known to reflect associative memory strength. For example, larger increases in A1 tonotopic area representing a sound frequency associated with reward correlates with stronger memory for that sound (Bieszczad & Weinberger 2010; Bieszczad et al., 2012). However, it is unclear whether subcortical plasticity in the auditory system parallels the cortical dynamics of learning-induced representational plasticity. The present study examined the relationship between memory strength and plasticity in lower auditory structures by using chronic auditory brainstem response (ABR) recordings over the course of two phases of auditory learning. Animals learned a tone-reward association to a high performance asymptote (>75% correct with 3 consecutive days of stable performance; C.V. ² 0.1) in order to produce a strong auditory associative memory. A subsequent extinction training phase was done to challenge the associative strength of the initial memory for sound. If lower auditory structures are sensitive to auditory associative memory strength, then tone-reward learning should induce plasticity detected in the ABR. Furthermore, the extinction challenge should oppose the ABR effect. In addition, the ABR evoked by the acoustic frequency of the signal tone should reveal greater representational plasticity than the ABR evoked by other frequencies if lower auditory structures are also sensitive to sound-specific memory. Indeed, the results showed that tone-reward training is associated with decreases in ABR peak latencies (relative to pre-‘tone-reward’ training latencies). While some of these ABR latency decreases endured extinction training, overall extinction was associated with increases in peak latencies (relative to post-‘tone-reward’ training latencies). Notably, peak latency changes exhibited some stimulus specificity: learning-induced ABR plasticity did not fully generalize to non-signal tone frequencies, particularly at frequencies more distant from the signal. Together, these data support the notion that, like the cortex, auditory brainstem plasticity is related to signal-specific memory strength. Moreover, the net change in the ABR after both tone-reward learning and its extinction reflected traces of both learning experiences, which suggests that the auditory brainstem is sensitive to the history of a sound’s significance. These results position auditory brainstem plasticity as a potentially useful predictive correlate of behavioral responses to meaningful sounds in animals with potential for application also in humans.
Unique roles for delta and theta frequency bands in the cortical analysis of temporal speech structure

Jackson C Lee (1), Aurelio J Falconi (1), Tobias Overath (1)

(1) Duke University

Speech perception entails the mapping of the acoustic waveform to its linguistic representation. For this mapping to succeed, the speech signal needs to be tracked across a large temporal range at high temporal precision in order to decode linguistic units (phonemes, syllables, words). Here we test how cortical processing of such temporal speech structure is modulated by higher-order linguistic analysis. To control the temporal scale of analysis, we used a novel sound-quilting algorithm that controls acoustic structure at different temporal scales (Overath et al., 2015). To control the linguistic content, we constructed speech quilts from both familiar and foreign languages. This ensures that any changes at the signal-acoustics level affect both languages identically, while manipulating the linguistic percept differently. Thus, neural responses that vary as a function of segment length, but are shared or similar across the two languages, suggest analysis at the signal-acoustics level, whereas neural responses that differ based on language familiarity imply the presence of linguistic processing. We recorded EEG while subjects listened to 6 s long English or Korean speech, quilted with 30 ms or 960 ms segment lengths. Neural entrainment to the speech quilt envelope, assessed via inter-trial correlation (ITC), in the theta band increased with segment length in both languages; however, ITC in the delta band increased with segment length only in English. This dissociation indicates that neural entrainment in the theta and delta frequency bands serves different functions: acoustic and linguist analysis of temporal speech structure, respectively. In particular, linguistic analysis tracks syllabic and word content that is preserved with increasing segment lengths. The results advance our understanding of the neural mechanisms underlying the acousto-linguistic mapping of temporal speech structure.
Cortically Inspired Spatial Processing Algorithm for solving the Cocktail Party Problem

Kenny Chou (1), Junzi Dong (1), H Steven Colburn (1), Kamal Sen (1)

(1) Boston University

The Brain’s ability to focus on a single acoustic stream within a complex auditory scene remains unmatched by machines, despite a long history of intensive research across various academic disciplines. At a cocktail party, we can broadly monitor the entire acoustic scene to detect important cues, or selectively listen to a conversation partner. This impressive dual-mode listening of normal-hearing listeners stands in sharp contrast with the difficulty faced by hearing impaired listeners, even with hearing assistive devices, and machine hearing algorithms. Recently, we developed a novel algorithm inspired by cortical neurons that display dual-mode responses to spatially distributed natural sounds. The algorithm processes binaural acoustic inputs using a cochlear filter-bank, a midbrain spatial localization network, and a cortical network. The output of the cortical network is then decoded into an acoustic waveform, using a stimulus reconstruction technique. Here, we compare the algorithm to state-of-the-art engineering algorithms to show the competitiveness of this algorithm. The cortically inspired algorithm may help improve the performance of hearing assistive devices, e.g., hearing aids and cochlear implants, and speech recognition technology currently challenged by cocktail party like settings; as well as computational algorithms for processing natural scenes cluttered with multiple objects.
Excitatory ON and OFF receptive field arrangement confers direction selectivity to slow frequency modulation in mouse auditory cortex

Gaelle Adeline CHAPUIS (1), Joseph SOLLINI (2), Claudia CLOPATH (1), Paul CHADDERTON (1)

(1) Imperial College London, London, United Kingdom
(2) UCL, London, United Kingdom

Frequency-modulation (FM) is an important feature in human and animal communication. In the auditory cortex, the relative tuning of excitatory and inhibitory synaptic input accounts for direction selectivity at high modulation rates (1). However, the mechanisms underlying direction selectivity at lower modulation rates relevant to vocalisation encoding have not been established. We used multi-electrode arrays to record the activity of mouse auditory cortex neurons to pure tones and ascending and descending logarithmic FM sweeps (range: 2.2 - 140 octave/sec). Single units in the auditory cortex of fentanyl-anaesthetised animals exhibited the strongest directional selectivity at low modulation rates (< 8 octaves/second). Surprisingly, we found that the relative tuning of excitatory ON and OFF responses to pure tones was an excellent predictor of the units' direction selectivity (correlation coefficient r = 0.61, p < 0.01, n = 30 cells with V-shaped tuning curves to onset and offset). These results suggest that cortically generated synaptic inhibition may not be required for direction selectivity at low modulation rates. To test this proposal, we have selectively expressed the inhibitory DREADD in parvalbumin-positive (PV+) interneurons to manipulate fast spiking interneurons within the auditory cortex (AAV1/2-hSyn-DIO-hM4Gi-mCherry injected into the auditory cortex of PV-Cre mice; n = 6). Following IP injection of the agonist, clozapine-N-oxide, we demonstrate a strong increase in local field potential response to sound, consistent with a local reduction in cortical inhibition. We are currently analysing the directional selectivity of single units and measuring behavioural discrimination to ascending and descending sweeps in mice trained to perform a Go/No-Go task under conditions of reduced excitability of fast spiking interneurons. These experiments will establish the relative roles of ON and OFF-mediated synaptic excitation and cortical inhibition in the representation of slow modulations in sound frequency.

1) Topography and synaptic shaping of direction selectivity in primary auditory cortex [doi:10.1038/nature01796]
Machine-learning classification of speech-evoked electroencephalographic signals reveals speech intelligibility

Youngmin Na (1), Subong Kim (2), Inyong Choi (2,3), Jihwan Woo (1)

(1) University of Ulsan
(2) University of Iowa
(3) University of Iowa Hospitals and Clinics

An important issue of hearing science is individual differences in the central processing for speech understanding. Although previous studies uncovered general cortical pathway for speech perception that involves the interplay between temporal, parietal and frontal lobes, how to quantify differential cortical processing among individual listeners is unclear. Here, we propose to use a machine learning-based classifier of speech-evoked electroencephalographic (EEG) signals to quantify fidelity of cortical speech processing. We hypothesize that important time-varying characteristics of acoustic and semantic features are extracted more accurately in listeners with better speech understanding and more strongly reflected in their cortical neural representation, which results in more accurate classification of time-sensitive EEG signals. To prove this concept, we compared the accuracy of EEG-signal classifier when it classified natural sentence-evoked potentials and vocoded sentence-evoked potentials. Given that a vocoder degrades speech intelligibility, we expected worse accuracy of EEG-signal classification when it is applied to the vocoded-sentence driven cortical signals.

We recorded 64-channel EEG signals from seven young normal hearing adults in response to ten natural sentences and their corresponding vocoded sentences in a passive listening condition. We used an eight frequency-band vocoder with continuous interleaved sampling strategy that mimics cochlear implant signal processing. Each sentence was repeated hundred times. One hundred random draws of ten trials (epochs) were used to compute event-related spectral perturbations (ERSPs). We divided ERSPs into five frequency ranges and four different time windows (40, 50, 100, and 200 ms), which produced 105-550 features (i.e. time-frequency bins) per epoch. We reduced the feature dimensionality using minimal redundancy maximal criterion. An artificial neural network with one hidden layer was trained with 80% of ERSPs. The other 20% of ERSPs were used for the cross-validation of classification accuracy.

Results confirmed our main hypothesis; classification accuracy was higher for natural sentence-driven EEG responses than vocoded sound-driven ones at frontal, central, and parietal electrodes in six out of seven subjects. This result demonstrates that a machine learning-based EEG classification can quantify the fidelity of speech comprehension. Now we are investigating whether this classification method is sensitive enough to reveal individual differences in normal hearing and hearing-impaired listeners’ speech understanding.
Spiking activity in auditory cortex to identity-preserving changes in sounds

Francisco A. Rodriguez Campos(1), Matthew Schaff(1), Josh H. McDermott(2), Yale E. Cohen(1)

(1) University of Pennsylvania

(2) Massachusetts Institute of Technology

The auditory system transforms acoustic stimuli into representations that subserve recognition and behavior. One major challenge is the creation of sound representations that are invariant to identity-preserving transformations. Indeed, sound recognition by human listeners is invariant to changes in the stimulus due to variation in the pitch of speakers’ voices, the location of the sound source, the amount of background noise, etc. The brain areas, neural mechanisms and computations that provide the basis for a listener’s tolerance to identity preserving changes are unknown.

To address the neural basis of invariant recognition, we conducted large-scale recordings from the auditory cortex of rhesus monkeys while they listened to natural sound exemplars (animal vocalizations and natural background noises) and identity-preserving transformations of these exemplars. These transformations included sound-source location and room reverberation. Additionally, as controls, we presented scrambled versions of these sounds. These scrambled versions were statistically matched to the natural exemplars but could not be identified as originating from the same source as the original sounds. We recorded with MicroProbes’ 96-electrode Microwire Brush Array simultaneously from different areas of auditory cortex.

We report how identity-preserving transformations of sounds are represented in the spiking activity of neurons in these three brain regions. In particular, we describe the degree to which neural populations codes support perceptual representations of invariance.
Developmental hearing loss impairs fast temporal processing

Justin D. Yao (1), Dan H. Sanes (1)

New York University

Fluctuations in sound level support speech comprehension, and perceptions of rhythm, prosody, and pitch, depending on the rate of amplitude modulation (AM). However, AM sensitivity may be vulnerable to hearing loss (HL) during its prolonged period of maturation. In fact, the processing of rapid (>100 Hz) modulation rates may be particularly impaired (e.g., Park et al., 2015). Here, we sought to determine whether behavioral and neural AM detection thresholds display a similar sensitivity to developmental HL. Normal hearing (NH) adult gerbils, and those reared with conductive HL, were trained and tested on an appetitive Go-Nogo AM detection task (Go stimuli were broadband noise modulated at 64-512 Hz across a range of depths; Nogo stimulus was unmodulated noise). HL animals displayed poorer behavioral thresholds than NH animals at all AM rates, but performance was much worse at very fast rates, 256 and 512 Hz. To determine whether these perceptual deficits were attributable to degraded encoding mechanisms within the auditory brainstem, we recorded envelope following responses (EFRs) across a range of AM rates and depths from NH and HL animals that had been trained previously on the AM detection task. NH animals displayed similar EFR and behavioral thresholds, whereas HL animals displayed EFR thresholds that were better than those obtained behaviorally. This suggests that a neural deficit associated with behavioral impairment may be found downstream of the auditory brainstem. To test this idea, we are recording extracellular responses from ACX neurons while gerbils perform the AM detection task.
58    Examining the Neural Representations of Auditory Objects: Speaker Identity and Musical Instrument Timbre

Mattson Ogg (1), Dustin Moraczewski (2), Stefanie E. Kuchinsky (3), L. Robert Slevc (4)
(1) University of Maryland, College Park
(2) University of Maryland, College Park
(3) University of Maryland, College Park
(4) University of Maryland, College Park

The ease and accuracy with which human listeners are able to identify sound sources in their environment belies the intricate neural processes that support such an ability. This perceptual feat is perhaps best exemplified by our ability to recognize individual speakers from the sound of their voice. Using pattern classifiers, it has been shown that the neural responses specific to individual human voices can be decoded from patterns of fMRI activation in the temporal lobes during perception. However, it is not clear whether other highly identifiable auditory objects relevant to humans, such as musical instrument timbres, can also be decoded in such a way. If possible, this would allow us to probe whether the brain codes representations of objects in a general manner (based on common neural substrates) or whether dedicated resources are devoted to separate categories of sound. Such analyses could also contribute to debates on the neural overlap between music and speech in which timbre or speaker identification, despite being an apparently salient parallel integral to perception in both domains, has received little attention.

To examine the representational specificity of auditory objects from different categories of sound, we presented 18 participants with a set of sound tokens consisting of 4 different speakers and 4 different instrument timbres while undergoing fMRI. We employed an interleaved silent steady-state acquisition sequence to maximize data collection while minimizing scanner noise to ensure that the participants accurately perceived the stimuli. We then decoded the pairwise combination of sounds within each category both in the whole brain and in a multi-voxel searchlight. The searchlight results yielded a cluster of voxels in anterior temporal cortex (bilaterally) that successfully decoded our set of individual speakers better than chance as well as a cluster of voxels in a slightly more posterior region of the temporal lobes that successfully decoded instruments better than chance. We also found that a small number of voxels were involved in the successful decoding of both categories of sound and were located in the temporal lobes between the posterior and anterior category-specific clusters. These results suggest that representations of auditory objects from different categories of sound engage at least partially overlapping neural resources in cortex, and this overlap exists perhaps as a function of localized clusters of neurons in the temporal lobes that represent auditory objects from individual categories.
Pressure difference receiving ears influence ITD coding in American alligators

Lutz Kettler (1), Catherine Carr (1)
(1) University of Maryland

The difference between the timing of sounds at both ears (intersaural time difference, ITD) is a key feature for sound source localization. However, ITD coding strategies may vary across taxa, head sizes, and frequencies. Two of the proposed mechanisms are a place code of neurons with different best ITDs within an iso-frequency band (Jeffress-model), and a population code that utilizes the slope of the rate-ITD curves of the entire neuron population over the physiological range, even while best ITDs lie outside the physiological range. A place code is less accurate with increasing noise, when head sizes are small, and for low frequencies where the peaks of ITD tuning curves become very broad. Thus, a rate-code may provide more robust encoding of ITD at low frequencies.

In archosaurs, the detection of ITD is assumed to be consistent with a Jeffress-model-like place code, which consists of coincidence detectors and delay lines. An anatomical structure that resembles the Jeffress model has been found in the nucleus laminaris (NL) of archosaurs, including alligators and birds. However, an additional factor, namely pressure difference receiving ears, may influence ITD detection in archosaurs. Alligators, like most birds, are sensitive to ITDs at low frequencies, where coupled ears allow transmission of sounds through the dorsal and ventral sinuses that connect both middle ears and create pressure difference receiving ears. Internal coupling increases the range of interaural time differences, and thus compensates for small head sizes. Thus, a place code might be efficient despite a low frequency hearing range.

We performed in vivo electrophysiological experiments with American alligators to investigate the roles of evolutionary history and anatomical constraints on optimal ITD coding. Like in chicken and owls, best ITDs were broadly distributed in alligators but often outside the physiological range. Unlike in mammals, the slopes of the tuning curves did not uniformly cross 0°s. Models of the physiological range that included pressure difference receivers predicted that the range of ITD was increased with decreasing frequency, and aligned with the limit of best ITDs given by the period of the stimulus tone (pi-limit). The increase in ITD range with frequency, thus, circumvented the constraint of best ITDs lying outside the physiological range.

Alligators and other archosaurs therefore may share a place code of ITD detection that co-evolved with internal coupling of ears. Since mammalian ears are not coupled, small early mammals may not have been able to compensate for their narrow ITD range, and may have evolved a population coding strategy.
Computational modeling approach to understand mechanisms for binaural pitch fusion

Yonghee Oh(1), Lina A. Reiss(1)

(1) Oregon Health & Science University

Recent studies suggest that hearing-impaired (HI) listeners differ from normal-hearing (NH) listeners in how they integrate sound between the two ears. In HI listeners, binaural pitch fusion is often increased so that sounds that differ greatly in pitch between the two ears are still heard as a single percept (Reiss et al., JARO 2014; JASA 2017; Oh and Reiss, JASA 2017). Furthermore, this broadened binaural fusion can cause binaural interference in HI listeners for vowel perception (Reiss et al., JARO 2016) and for speech understanding in background noise (Oh et al., ASA 2017).

The goal of this study was to investigate how within-ear auditory sensitivity may affect binaural pitch fusion using experimental and modeling approaches. Psychophysical measures of within-ear frequency tuning, including audiograms and psychoacoustical tuning curves, were obtained and characterized as model parameters for both NH and HI listeners. The model performance was then evaluated on binaural pitch fusion measurements. Results show that broad fusion in HI listeners was positively correlated with broadly tuned auditory filters. In addition, the model output predicted the variation in individual binaural fusion data in both subject groups.

These findings suggest that monaural frequency tuning, which may reflect peripheral sensitivity, could be a potential factor related to the non-optimized hearing mechanisms involved in the abnormal binaural auditory processing for HI listeners. Such model-based interpretation could indicate the relative importance of peripheral and central substrates in degraded benefits of binaural processing for speech perception.

Supported by NIH/NIDCD R01 DC01337, P30 DC005983, and F32 DC016193.
Auditory fear conditioning drives changes in frequency representation and functional organization of neuronal populations in the auditory cortex

Katherine C Wood (1), Richard Betzel (1), Christopher Angeloni (1), Danielle Bassett (1), Maria N. Geffen (1)

(1) University of Pennsylvania

Everyday auditory behavior depends critically on learning-driven changes in auditory perception that rely on neuronal plasticity within the auditory pathway. Discriminative auditory fear conditioning (DAFC), a form of associative auditory learning, affects the fundamental auditory task of frequency discrimination acuity. Previous work has shown that the auditory cortex (AC) is required for this modulation and that DAFC induces changes in individual frequency tuning of cortical neurons, with the best frequency shifting toward the paired conditioned tone (CS+). However, how learning shapes frequency representation at the level of neuronal populations remains unknown. To understand the transformation in tone representation in AC before and after conditioning, we recorded activity of populations of hundreds of neurons simultaneously in AC of awake, head-fixed mice, tracking the same neurons over days, using two-photon imaging of Calcium activity. We then quantified changes in tone frequency-dependent responses and functional connectivity structure of these populations.

We found that over successive days, even prior to learning, tone-evoked responses of individual neurons were variable, with best frequency of neuronal tuning shifting from day to day. However, when averaged over the neuronal population, the representation of tones within specific frequency bands was preserved. We then used a novel method, based on graph theory, to demonstrate that over the neuronal population, the spontaneous activity exhibited a clustered correlational structure. Specific clusters contained more tone responsive neurons than would be expected by chance suggesting an underlying functional structure.

Following learning, individual neurons exhibited shifts in best frequency both toward and away from CS+ or the unpaired CS-. However, over the neuronal population, there was an increase in activity in response to both CS+ and CS-. These changes in the response amplitude at CS+ and CS- were significantly correlated. After fear conditioning, the cluster structure of the network architecture was preserved.

Our study demonstrates that auditory fear conditioning drives changes in cortical neuronal populations both at local and global level, leading to enhanced representation of both the paired and the unpaired conditioned tones.
HDAC3-inhibition facilitates long-term memory consolidation of sound-specific information

Andrea Shang (1), Sooraz Bylipudi (1), Kasia M Bieszczad (1)

(1) Rutgers University -- New Brunswick

Neuroplasticity and successful long-term memory (LTM) can be modulated by epigenetic mechanisms, such as chromatin modification. Histone deacetylases (HDACs) typically oppose histone acetyltransferases (HATs) and repress gene expression by constricting chromatin. HDAC inhibitors have therefore been shown to open chromatin and facilitate various forms of LTM (e.g. McQuown & Wood, 2011; Malvaez et al., 2013). For example, rats treated with a pharmacological HDAC3-inhibitor (HDAC3i) during consolidation of an auditory associative task have enhanced memory for the behaviorally-salient sound-frequencies in a memory test after training, as well as expanded representation in primary auditory cortex (A1) for the signal-frequencies (determined by electrophysiological 'mapping' of A1 tonotopy after all training and testing; Bieszczad et al., 2015). Here, we test the emerging hypothesis that HDAC3 controls the sensory-specificity of LTM (Phan & Bieszczad, 2016). Adult male rats (N=45; Sprague-Dawley) were trained on a two-tone frequency-discrimination (2TD) task with immediate post-training injections of an HDAC3i (RFGP966, Abcam Inc.; 10mg/kg, s.c.) or vehicle (Veh) for 3 consecutive days. Rats learned to associate one of two spectrally-different pure tones with reward over ~3 weeks (5.0 kHz, CS+; 11.5 kHz, CS-). HDAC3i facilitated the rate of 2TD task acquisition: treated rats (N=12) reached performance criterion faster than vehicle controls (N=12). Furthermore, after both groups reached the same level of asymptotic performance, a stimulus generalization memory test (SGT) was administered. BP responses elicited by different sound frequencies (including the CS+ & CS-) were recorded under extinction conditions. Groups differed in the proportion and latency of BPs to SGT test tones, which indicated that HDAC3i-treated rats (N=9) had formed more specific memory for both CS frequencies compared to Veh-treated rats (N=9). A subset of animals also underwent a 2-week delayed SGT (HDAC3i, N=4; Veh, N=4) and were immediately re-tested for 2TD performance. Resilient performance in HDAC3i-treated rats weeks after HDAC3i injections suggest that limited treatment during initial task acquisition has lasting effects on frequency-specific memory. A1 recordings after immediate and 2-week delayed SGTs reveal how tonotopic representation of different associative sound-signals (CS+ and CS- frequencies) interacts in animals that have specific vs. non-specific memories. These data support that HDAC3 is key negative regulator of the sensory specificity of LTM formed.
Robust discrimination of sounds embedded in noise by adapting cortical gain

Chris F. Angeloni (1,2), Katherine C. Wood (2), Maria N. Geffen (1,2)

(1) Psychology Graduate Group, University of Pennsylvania
(2) Department of Otorhinolaryngology, University of Pennsylvania

Natural acoustic environments are variable; to function in them it is necessary to reliably differentiate important sounds from one another despite the presence of a noisy background. Recent work demonstrated that neurons in auditory cortex modulate their response gain in an adaptive manner to account for changes in the spectrotemporal statistics of acoustic stimuli (Rabinowitz et al., Neuron, 2011; Natan et al., Cereb. Cortex, 2016). In the current study, we model how such changes in neural gain facilitate robust discrimination of stimuli embedded in different types of noise. Using a linear-nonlinear probabilistic model of neural activity, we simulated neural responses to sounds embedded in noise backgrounds of different contrast levels. Then, using techniques from signal detection theory, we estimate behavioral detection and discrimination from the predicted neural responses and how behavioral performance changes depending on the gain of the neuron and the loudness of the embedded signals relative to the mean noise level (i.e., their signal to noise ratio, SNR). We find that detection of a signal embedded in noise is always facilitated by high gain, but does not change appreciably as the neuron adapts to a new noise context. However, discrimination in different noise contrasts depends on the gain state of the neuron when differentiating between targets embedded at different SNR levels. Namely, target stimuli with high SNR are readily discriminable in high contrast noise, but not low contrast, while target stimuli with low SNR are more discriminable in low contrast than high contrast. In this framework, we also observe differences in the time course of discrimination performance, such that during high to low contrast adaptation, discriminability decreased earlier and faster for higher SNR targets, while the opposite effect is seen in low to high contrast. These findings provide a predictive framework in which we can model detection and discrimination tasks in the presence of a noise background, the subsequent neural responses to target stimuli as a function of gain adaptation, and how those neural responses predict behavioral performance.
Predicting the neural responses to speech in human auditory cortex using deep neural network models

Hassan Akbari (1), Bahar Khalighinejad (1), Jose L. Herrero (2), Ashesh. D. Mehta (2), Nima Mesgarani (1)

(1) Columbia University
(2) Hofstra Northwell School of Medicine

Recently, interest has grown in characterizing the response properties of sensory neurons under natural stimulus conditions. The majority of previous studies have used linear models to relate the acoustic features of sound to neural responses. However, linear models cannot capture the inherent non-linearity of the responses in the brain. We investigated the utility of deep neural network models to predict neural responses to speech in human auditory cortex. The neural responses were recorded from the transverse and the superior temporal gyrus of patients undergoing surgery for the treatment of epilepsy as they listened to continuous speech. Deep neural networks (DNNs) have shown great promise in capturing non-linear relationships. Thus, we trained a DNN with a non-linearity in each layer using the time-frequency representation of the stimulus as the input and the envelope of the high-gamma activity of the neural responses as the output of the model. First, we started with a one-node one-layer fully connected network, which is equivalent to the commonly used spectrotemporal receptive field (STRF). We then proceeded by progressively adding nodes and layers to study the effect of deepness and complexity of the model on the accuracy of predictions. In addition, we explored the effect of network architecture such as convolutional (CNN) and recurrent (RNN) neural networks.

To compare the prediction of networks (STRFs, DNNs, RNN, and CNNs), we obtained the difference between the predicted responses of the DNNs and the STRF model and determined the response features that were better predicted by the nonlinear models. We also analyzed the network to figure out the properties of the functions that are applied to the stimulus. In comparison to the STRFs, the predicted responses from the neural networks had a higher correlation with the original responses. On average, using DNNs improved the correlation by 5%, and the CNNs by 25%. We found that a combination of DNNs and CNNs can provide a 35% improvement, which suggests that more complicated networks are needed to model the neural encoding of speech. We further analyzed the properties of the functions learned by the networks to shed light on the computational role of the nonlinear transformations that are applied to the sound features as they travel throughout the auditory pathway.
Somatosensory-auditory interactions in auditory thalamus

Michael Lohse (1), Johannes Dahmen (1), Victoria M. Bajo (1), Edward Mann (1), Andrew J. King (1)

Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, United Kingdom

Considerable attention has focused on the involvement of auditory cortex in multisensory processing. Although visual and somatosensory influences can arise via corticocortical pathways, it is likely that some aspects of multisensory cortical processing are inherited from the thalamus. However, we currently know much less about the origins and nature of visual and somatosensory influences on the activity of neurons in the auditory thalamus. We investigated multisensory interactions (somatosensory-auditory) across different divisions of the medial geniculate body (MGB) and the circuits underlying these multisensory effects.

We found that deflection of all principal whiskers differentially modulates responses in the auditory thalamus along a latero-medial gradient. Auditory responses were found to be reliably suppressed in both the ventral and dorsal divisions of the MGB. Auditory responses in the more medial structures (medial division of MGB/posterior intralaminar nucleus and suprageniculate nucleus) were facilitated by somatosensory stimulation.

Somatosensory-induced influences were also found in auditory cortex, but absent in the central nucleus of the inferior colliculus. We found that these effects could to a large extent be mimicked by channelrhodopsin-2 stimulation of neurons in corticofugal layers of primary somatosensory cortex. This suggests a crossmodal corticothalamic origin of the multisensory modulations found in auditory thalamus. We are furthermore investigating the involvement of the thalamic reticular nucleus in mediating the somatosensory suppression of auditory responses in the ventral and dorsal divisions of MGB.

Our results suggest that the sensory thalamus, including the principal relay nucleus, processes multisensory information, and that these interactions are mediated by corticothalamic circuits.
Neuronal activity in the human subthalamic nucleus during speech production

Araceli Ramirez-Cardenas (1), Kris Tjaden (2), Lisa Kopf (3), Haiming Chen (1), Karen N. Bryant (3), Daniel M. Corcos (4), Jeremy Greenlee (1)

(1) Univ. of Iowa Hosp. and Clinics, Dept. of Neurosurgery.
(2) University of Buffalo
(3) University of Iowa
(4) Northwestern University, Physical Therapy and Human Movement Sci.

Like other motor behaviors, speech requires the sequencing and control of simpler movements into a coordinated action. Accordingly, the basal ganglia are thought to play an important role in speech programming, initiation, execution, and control. Moreover, recent evidence suggests that basal ganglia involvement extends beyond the strict motor aspects of speech, to a role in language and decision making. This is particularly the case for the subthalamic nucleus (STN). In the present study, we want to investigate whether neural activity in the subthalamic nucleus is modulated by speech complexity. For that purpose, we recorded LFPs and spiking activity in the subthalamic nucleus of four patients who underwent deep brain stimulation (DBS) surgery for clinical reasons. Data were collected while subjects produced utterances with different levels of complexity (sustained vowels, word repetition, sentence repetition and free monologue). Subject utterances were captured with a microphone for further examination. Preliminary analyses of the data show a modulation in beta power for all types of utterances. Particularly, beta power was suppressed during speech production. Our results advance our understanding on the role of the basal ganglia in language processing and production.
Comparing Human and Nonhuman Primate Brain Responses to Auditory Sequences Using EEG

Daniel J Cameron (1), Luis Prado (2), Jessica A Grahn (3), Hugo Merchant (2)

(1) Georgetown University
(2) Universidad Autonoma Nacional de Mexico
(3) University of Western Ontario

Humans have the ability to perceive and move to the ‘beat’ in musical rhythms. This complex behaviour may be unique among primates, and involves extracting regularities from complex, non-isochronous auditory sequences and synchronizing movements to those regularities. To characterize cross-species differences in neural processing of rhythm and beat perception, we compared brain responses in humans and a macaque monkey using electroencephalography (EEG). EEG was recorded while participants listened passively to auditory sequences that consisted of white noise bursts (65ms) separated by intervals. Intervals were structured to create 5 types of sequences: i) isochronous (at three rates), ii) random, iii) strongly beat-based, iv) weakly beat-based, or v) non-beat-based. We compared the two species’ evoked and induced responses in EEG, at various frequency bands (e.g., beta and gamma) to isochronous and random sequences, and compared neural entrainment at low-frequency (1-5 Hz) regularities in the structures of non-isochronous sequences (iii-v). Similarities and differences between species’ brain responses reveal common and distinct aspects of neural processing of auditory sequences.
Humans have the remarkable ability to selectively focus on a single speaker in a multi-speaker environment. The neural mechanisms that underlie this phenomenon nonetheless remain incompletely understood. In particular, there has been longstanding debate over whether attention to speech operates at an early stage in the speech processing hierarchy based on physical characteristics of the stimulus or later on during higher-order linguistic processing. Progress to resolve this has been hindered by the fact that fMRI lacks temporal resolution, whilst ERP-based EEG studies limit paradigms to the use of discrete stimuli such as isolated syllables and words.

Recently, the use of system identification to linearly fit models that map between stimulus features to concurrently recorded EEG has been found to be effective for studying the processing of natural, continuous speech (Lalor & Foxe, 2010) and cocktail party attention in particular (Power et al., 2012). Using this method, EEG has been found to be sensitive to both low-level acoustics as well as higher-level categorical speech features like phonemes (Di Liberto et al., 2015).

Here, we employ this approach in a multi-speaker paradigm to identify differences between phoneme-level processing in attended and unattended speech. Participants attended to one of two concurrently presented speakers (perceived to be positioned on their right and left) whilst scalp EEG was recorded. We computed models -- known as temporal response functions -- that map from the attended and unattended phoneme-level features to EEG. We then assessed these models in terms of their ability to predict novel EEG data during cocktail party attention. Comparison of model performances for attended and unattended speech permit the assessment of how auditory selective attention affects the processing of phonemes, a high-level speech characteristic. These effects are compared to those found when relying on lower-level speech representations like the envelope.
Verbal communication in noisy environments is possible because of selective attention, a cognitive enhancement of attended stimuli and suppression of distractors. However, individuals differ in their ability to parse out, monitor, and sustain multiple streams of simultaneous stimuli due to individual differences in top-down, executive function. Here we have developed analogous auditory and visual tasks to probe how individuals’ top-down functioning modulates neural correlates of selective and divided attention across modalities using electroencephalography (EEG).

In the auditory task, human subjects listened to three time-staggered, spatially lateralized streams of speech reciting permutations of the syllables bah, dah, and gah. Subjects were prompted to identify the order of the syllables presented from the center ‘target’ stream or the left ‘supertarget’ stream and to always ignore the right ‘distractor’ stream. In a random subset of trials, the supertarget did not appear, and subjects instead reported the target stream. This setup resulted in two attentional states: focused attention, where only one stream is attended, and broad attention, where two streams must be monitored to prepare for a spatial attention switch. The corresponding visual task flashed the letters B, D, and G on a blank screen at three different positions (left, center, and right) with onsets identical to those of the auditory task’s stimuli. Subjects fixated their gaze on a stationary dot and employed covert visual attention to identify the order of the letters in the prompted streams. Scalp potentials were recorded using a 64-channel EEG setup.

In both tasks, event related potential N1 amplitudes evoked by target stimuli were smaller in broad attention trials than in focused attention trials and, accordingly, performance was lower in the broad condition - particularly when no supertarget appeared - implying that across modalities, dividing attention introduces performance costs. Individuals’ N1 amplitude ratio, a measure of N1 enhancement to supertarget relative to target stimuli, correlated with performance in both modalities. Additionally, in the auditory task, oscillations in the alpha band (8-12 Hz) were increased in parietal-occipital areas contralateral to the suppressed supertarget stream in the focused condition, implying that the spatial nature of the task engages visuospatial networks for both modalities. These results suggest the need to further study individual differences in selective attention modulation across sensory modalities, particularly as executive attentional disorders such as ADHD and traumatic brain injuries are increasingly diagnosed.
Assessing musical beat perception based on simulated low-level neural activity

Nathaniel Zuk (1), Edmund Lalor (1 & 2)

(1) University of Rochester
(2) Trinity College

Prior research has shown that musical beats are salient at the level of the auditory cortex, as measured by recording cortical activity in humans. Several models have been developed that infer beat timing from stimuli with discrete temporal events, usually clicks or pure tones, which simulate our perception of beats and the observed neural activity in the cortex. Yet below the cortex there is a considerable amount of low-level processing that could influence our perception of beats. How such low-level processing shapes beat perception, especially in real music, is still unclear.

Here, we use a model of the auditory nerve [Zilany et al, J Acoust Soc Am, 135, 283-286 (2014)] and a model of midbrain-level modulation filtering [Carney et al, eNeuro, 2 (2015)] to simulate low-level neural activity with audio frequencies from 125 Hz to 8 kHz. We then examined how the beat frequency can be determined based on this neural activity using 1163 segments of music, each 20-30 seconds in duration, that have previously been used for assessing tempo identification algorithms. We show that the most synchronous neural activity summed across frequency often occurs at the musical ‘tactus’ or event frequency rather than at the beat frequency. Then, we show that by dividing the event frequency by a scaling factor classified using the event frequency and information about the music’s rhythm, tempo identification performance improves. Our result suggests that detecting regularities in bottom-up responses is not sufficient for perceiving a beat in real music; an additional, potentially cortical-level, mechanism is necessary.
The effect of musical experience on the organization of neural stimulus selectivity in human auditory cortex

Dana Boebinger (1,2), Josh McDermott (1,2), Nancy Kanwisher (2)

(1) Speech and Hearing Bioscience and Technology, Harvard University

(2) Brain and Cognitive Sciences, MIT

Speech and music perception are core functions of the human auditory system, but we know little about the neural systems that process these complex auditory signals. While speech-selectivity is known to characterize regions of non-primary auditory cortex, clear neural selectivity for music has only recently been demonstrated (Norman-Haignere, Kanwisher, & McDermott, 2015). This study measured fMRI responses to a wide variety of natural sounds and modeled the response of each voxel as a weighted sum of canonical response patterns to the sound set (‘components’). Six components explained 80% of the replicable variance in the responses across auditory cortex. Four of these components reflected selectivities for acoustic features of the sounds (e.g. frequency, spectrotemporal modulation), and two showed selectivity for the high-level categories of music and speech. When these speech- and music-selective components were projected back onto the brain, they concentrated in non-overlapping regions of non-primary auditory cortex, and their anatomical distribution was similar across subjects.

This new method for parcellating neural stimulus selectivity allows us to ask many questions. Here, we replicate the previous finding of speech- and music-selective components, and we examine the extent to which this component structure is altered by experience. Specifically, we report the extent to which musical training affects stimulus selectivity in auditory cortex by comparing the anatomical distribution and selectivity of response components in expert musicians versus control participants with minimal musical training.
Primate vocal communication is characterized by the production and perception of vocalizations within the context of active social behaviors. Within this complex system, what is communicated by a vocalization is heavily dependent on the immediate social context in which it occurs. However, much of what is known about the neural basis of primate communication comes from traditional studies involving subjects that are either passively listening or trained to respond to vocalizations with a conditioned behavioral response. Because of the active nature of communication, each of these more conventional approaches divorces the signal from the core context in which it naturally occurs. Here we sought to test this critical issue by recording the responses of frontal cortex neurons in marmosets across multiple behavioral contexts. Specifically, we presented the same vocalizations (phee calls, twitters) and a variety of white noise stimuli to marmosets while they were head-fixed and freely-moving using identical stimulus presentation protocol. During each of these test sessions, we also recorded the same neurons while subjects engaged in their natural antiphonal calling behavior, which involves the reciprocal exchange of phee calls. Preliminary analyses revealed that (1) white noise elicited the most consistent and robust responses across all test stimuli, (2) only half of neurons responsive to an acoustic stimulus in the head-fixed condition exhibited a similar response to the same stimulus in the freely-moving condition and (3) phee calls elicited the weakest responses of all stimuli in all conditions. These findings suggest that neural responses under head-fixed, passively listening conditions may not be predictive of more natural contexts. We are currently examining a broader set of stimuli in order to test the hypothesis that distinct neural processes may underlie vocal signal processing during natural, active communication and these other behavioral contexts.
Evolution of vocal circuits; analyses of hybrid song features in crosses between Xenopus laevis and Xenopus petersii

E. PEREZ (1,2), C. L. BARKAN (3), I. C. HALL (4), J. SEGARRA (1), S. M. WOOLLEY (2), D. B. KELLEY (1)

(1) Dept of Biological Sciences, Columbia University, NY
(2) Psychology, Columbia University, New York, NY
(3) Reed college, Portland, OR;
(4) Biological Sciences, Benedictine University, Lisle, IL

African clawed frogs (Xenopus) are nocturnal and aquatic anurans; vocal signaling dominates social communication. During courtship, males produce a distinctive advertisement call with species-specific temporal and spectral features. These calls consist of trains of sound pulses (trills); each pulse consists of two simultaneous, dominant frequencies (DFs). Within the laevis clade, X. laevis and X. petersii calls include fast and slow trill components with distinct DF2/DF1 ratios. Female Xenopus are more responsive to conspecific than to heterospecific calls: specifically sensitive to the spectral features of their own species’ call, a difference that could contribute to premating isolation and species divergence. We recorded advertisement calls from sexually mature X. laevis/X. petersii hybrids and analyzed the temporal and spectral features of their advertisement calls. Call duration, call period, fast trill duration and slow trill duration in male hybrids differ from both parental species but are closer to X. petersii than X. laevis while DF2, DF1 and DF2/DF1 are intermediate. To assess the spectral sensitivity of female hybrids, we are comparing auditory-evoked potentials (AEPs) elicited by synthetic two-tone sound pulses (dyads) - characteristic of both parental species and of hybrids - to shifted dyads. Ancient interspecific hybridizations have shaped the phylogenetic relationships of extant Xenopus. Divergence between X. laevis and X. petersii in the duration of fast trill and call period are controlled by differences in intrinsic rhythmic features of vocal pattern generator within the parabrachial area of the rostral hindbrain. Both male and female hybrid offspring are inter-fertile, facilitating the generation of backcrosses and intercrosses in which vocal features and auditory sensitivity can be mapped to genomic loci. Thus analysis of vocal patterns in male and spectral sensitivity in female X. laevis/X. petersii hybrids may provide insights into mechanisms that drive the divergence of new species.
Cortical mechanisms underlying responses to temporal gaps

Bshara Awwad, Israel Nelken

Dept. of Neurobiology, Institute of Life Sciences, and the Edmond and Lily Safra Center for Brain Sciences, Hebrew University, Jerusalem, Israel.

High temporal acuity of auditory processing is crucial for speech understanding. A Gap-detection test is used clinically to measure auditory temporal acuity. Abnormal gap detection thresholds are believed to reflect central hearing difficulties even in the presence of normal hearing sensitivity. In spite of their clinical relevance, cortical mechanisms for gap processing are poorly understood.

Using an oddball paradigm, we studied the coding of gaps in the rat auditory cortex. Gaps are composed of two noise markers separated by the gap duration. The sensory response to a gap consists of two components (called here P1 and P2), representing the responses evoked by the two markers. In cortex, these are usually mostly tonic responses. Because of forward masking, the P2 response is smaller than the P1 response, so that the P2/P1 ratio is typically smaller than 1.

We recorded responses to sequences consisting of either gap stimuli or continuous noise stimuli, and manipulated the probability of gap stimuli in the sequences. When the gap stimuli were rare (gap deviants), the P2/P1 ratio was greater than when the gap stimuli were common (gap standards). Moreover, in many neurons, when the gap was long (20ms), the P2/P1 was greater than 1, reversing the ubiquitous effect of forward masking. Thus, neurons in rat auditory cortex show strong Stimulus Specific Adaptation (SSA) to gaps, even at detection threshold (~ gap duration of 2ms).

Using intracellular recordings, we observed that in the P2 the excitation-inhibition (E/I) balance of the sensory response changes, with excitation becoming more prominent, accounting for the larger response of P2 when the gap is deviant. We hypothesize the existence of a class of neurons that are activated only by the second marker of the gap, and who are responsible on the large P2 excitatory response; such responses may arise from the action of disinhibitory circuits that are known to be present in cortex.
Interaction between top-down and bottom-up predictions in auditory decision-making

Lalitta Suriya-Arunroj (1,2), Yale E. Cohen (1,2,3), Josh I. Gold (2)

(1) Department of Otorhinolaryngology-Head and Neck Surgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
(2) Department of Neuroscience, University of Pennsylvania, Philadelphia, PA, USA
(3) Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA

Auditory perception and decision-making are mediated by bottom-up and top-down processes. Although a perceptual decision is based primarily on incoming sensory stimuli (bottom-up processing), it can be biased by different cognitive and task-relevant pieces of information. For example, prior knowledge and attention can facilitate and modulate this decision process by prioritizing relevant stimuli or features of stimuli (top-down processing). In contrast, changes in stimulus salience, such as regularity violation, can also bias perceptual decision-making through bottom-up attention. However, the effects that top-down and bottom-up processing have on auditory decision-making are not well understood. Here, we investigate how top-down and bottom-up predictions influence auditory decision-making.

Human subjects listened to sequences of high-frequency (2000 Hz) and low-frequency (500 Hz) tone bursts (tone duration: 300 ms with a 10-ms cos2 ramp; inter-burst interval: 100 ms) embedded in background, broadband noise and played through headphones. Subjects reported whether the last tone in each sequence (i.e., the ‘test tone’) was ‘low-frequency’ or ‘high-frequency’ by pressing the left or right button, respectively, on a gamepad. We titrated difficulty by varying the amplitude of the test tone relative to the noise masker. Top-down processing was manipulated by presenting a visual pre-cue that indicated the prior probability that the test tone would be high or low frequency for the given block of trials (corresponding to ratios of low- versus high-frequency tones of 5:1, 1:1, 1:5). Bottom-up processing was manipulated by varying the sequence of tone bursts presented just before the test tone ('pre-tones'), in terms of: 1) the ratio of high- and low-frequency pre-tones, which was either highly regular (all high or all low) or irregular (randomized); and 2) the number of pre-tones, which were selected from an approximately exponential distribution (range: 2--10) to minimize the ability to predict the timing of the test tone.

Preliminary results indicate that both the top-down and bottom-up manipulations caused choices biases. These biases were largest when stimulus discriminability was lowest, consistent with principles of signal detection theory. These biases also tended to be larger in response to the bottom-up manipulations. Further analyses of choice and response times will help identify the computations used to incorporate these two sources of information into perceptual decisions.
Differential encoding of auditory information by subgroups of neuronal spikes associated with coordinated neuronal ensembles

Jermyn See (1), Craig Atencio (1), Vikaas Sohal (1), Christoph Schreiner (1)

(1) University of California, San Francisco

There is mounting evidence suggesting that sensory stimuli are processed by interconnected populations of neurons. However, most studies of information processing in the primary auditory cortex (AI) involve either single-unit spectrotemporal receptive field (STRF) estimation or paired neuronal correlation analyses, and assume that AI neurons filter auditory information either as individual entities or as pairs. Determining how AI encodes information will require an integrated approach that combines receptive field and multi-neuronal ensemble analyses.

To assess multi-neuronal information processing in AI, we performed dense extracellular recordings in rat AI while presenting dynamic, broadband stimuli. We used dimensionality reduction techniques to identify distinct groups of AI neurons (coordinated neuronal ensembles, or cNEs) that have reliable synchronous activity. We identified cNE events and used them to assess spectrotemporal information processing. Neuronal spikes associated with cNE activity conveyed greater information than spikes that were not associated with cNE activity. For neurons that participated in multiple cNEs, spikes associated with one cNE had receptive field properties that were significantly different from that of spikes associated with other cNEs.

These findings challenge the classical idea that AI neurons produce a homogeneous set of spikes that may be equally weighted to estimate a single STRF. Instead, AI neurons can have multiple receptive fields based on associations with different cNEs, with each cNE representing the convergence of thalamocortical, intracortical and top-down inputs into AI. For each AI neuron, equally weighting all neuronal spikes to form a single STRF ignores this enhanced coding capacity. Therefore, by taking into account the stimulus preferences associated with each cNE, we may gain a more complete evaluation of information processing in AI.
Reliable and persistent population encoding of sounds across layers in the auditory cortex of awake mice

Daniel E. Winkowski (1), Zac Bowen (1), Tiago Ribeiro (2), Dietmar Plenz (2), Patrick O. Kanold (1)

(1) Dept of Biology, Univ. of Maryland, College Park, MD;
(2) Section on Critical Brain Dynamics LSN/NIMH, NIH, Bethesda, MD;

The ability to detect and encode sensory stimuli hinges on the encoding capability of neuronal populations. Individual neurons in primary auditory cortex (A1) can show selectivity to stimulus features (i.e., sound frequency). However, the representation of sound information in populations of A1 neurons and the nature of populations activated by sound stimuli is unresolved. The theory of critical dynamics describes complex systems that operate at or near a critical point, i.e., balanced between order and disorder, where systems gain advantages in information processing and their dynamics become scale-invariant. This theory has been gaining traction in neuroscience since the identification of neuronal avalanches, intermittent spatiotemporal activity patterns with scale-invariant properties, e.g. their size and duration, are power law distributed.

Here, we tested whether sound processing in A1 neuronal populations of Layer 2/3 (L2/3) and Layer 4 (L4) is in line with expectations for critical dynamics. We characterized spontaneous as well as sound evoked activity in L2/3 and L4 in A1 of awake mice using in vivo 2-photon calcium imaging of neuronal populations. We explore the relationship between properties of sensory stimuli, single neuron responses, and population responses with respect to avalanches and compare our findings across cortical laminae. We investigate the spectral tuning properties of neurons participating in avalanches in order to probe the relationship between criticality and sensory coding. On the population level, we found that avalanche statistics varied with stimulus frequency and intensity of the presented sound. In L2/3, we found that avalanche patterns reliably represented sounds independent of level and contained stimulus information well after sound offset, revealing long-range spatiotemporal activity correlations not evident at single-cell or cell-pair level. Ongoing work is investigating whether this persistent encoding of stimulus information is inherited from L4 populations or whether this is an emergent feature of population activity in L2/3.

Our investigation will provide insight into how neural networks containing differing populations of neurons with varying firing rates stably encode information about sensory stimuli in the context of critical dynamics.
Vocal communication plays a key role in maintaining group cohesion and social bonds for group-living primates. Successful vocal communication requires signalers to inform receivers about the social context of the produced vocal signals (e.g., vocal signals are directed toward specific receivers). Marmoset monkeys (Callithrix jacchus) exhibit many human-like social behaviors, including interactive vocal communication. It is unclear, however, the degree to which the natural variability of marmoset vocalizations is deliberate and informative versus a result of random fluctuation. Previous experiments using isolated pairs of marmosets have shown that marmosets produced acoustically different phee calls during vocal interaction than during spontaneous calling. However, vocal communication in naturalistic social environments involves many types of vocalizations and many group members. Therefore, it remains unknown if marmosets produce acoustically different calls depending on details of the social context, and whether such vocal acoustics influence the responsiveness of others. We recorded vocalizations from an entire group of marmosets housed in a colony environment and measured their vocal interactions in order to determine which animals responded to other’s vocalizations and in what order. We particularly focused on the marmoset trill call, frequently produced in the housing environment, and found that acoustical properties (e.g., duration) of a produced vocalization influenced the probability of evoking a vocal response. Furthermore, such response probabilities differed between different producer-responder pairs, suggesting an association between the acoustic properties of produced vocalizations and their intended receivers. Additionally, we further tested whether marmosets are sensitive to vocal acoustic variability by playback of computer-controlled vocal sounds in the housing environment. We found that playback of vocalizations evoked responsive vocal production from different individuals within colony. Together, these results further elaborate our understanding of the complex social mechanisms used by marmosets during vocal communication that can potentially be exploited to understand the neural basis of social behaviors.
Stimulus dependent versus cell specific computational strategies in localization of two simultaneous sounds

Shawn M. Willett (1), Valeria C. Caruso (1), Surya T. Tokdar (1), Jennifer M. Groh (1)

(1) Duke University

Coding of multiple simultaneous sounds is a known problem for broadly tuned neural representations of auditory space. Previous reports suggested neurons in monkey inferior colliculus (IC) implement time-division multiplexing, or alternation between firing rates corresponding to each stimulus (Caruso et. al., BioRxiv, 2017; Willett et. al., Soc Neuro Abstr 2016). These studies found a variety of fluctuating and stable response patterns across the set of tested conditions. Here, we investigated whether the same neurons tested with different stimuli tended to demonstrate the same vs. different types of response patterns.

Single unit activity was recorded in the IC of a rhesus macaque performing single- and dual-sound localization tasks involving one or two saccades to the location of each stimulus. Spike counts on dual sound trials were modeled in relation to the Poisson distributions of spike counts observed on single sound trials. Responses were classified as deriving from (a) a mixture of the two single-sound Poissons, (b) an intermediate Poisson distribution with a rate between those of the two single sounds, or a Poisson distribution that (c) matched or (d) was outside the range of the single-sound Poissons. We then determined whether the winning model was consistent for all conditions an individual neuron was tested with.

The results were heterogeneous. Roughly 45% of cells appear to use a cell specific computation, with the majority of winning models being consistent across the tested conditions. Another 37% of cells exhibited distinctly different behavior across different stimulus conditions. Together, these results indicate that IC neurons can either play a consistent type of role in the population response to dual sounds or they can be flexible and respond differently to different combinations of stimuli.
Activation and deactivation of somatostatin-expressing GABA neurons disrupts temporal coding of sound frequency in the auditory cortex

Andrew T. Landau (1,2), Wei Guo (3), Daniel B. Polley (1,2)

(1) Harvard University
(2) Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary
(3) Massachusetts Institute of Technology

The dynamic interplay between excitation and inhibition sculpts the neural representation of sound features at every stage of central auditory processing. At the level of the auditory cortex, where all inhibition is local, fast-spiking GABA cells balance the tuning of excitatory cells, but other populations of cortical GABA cells may play a crucial role in shaping cortical responses to sound. Here, we use an optogenetic strategy to manipulate and record from dendrite-targeting somatostatin-expressing (SST) GABA neurons in the primary auditory cortex (A1) of awake, head-fixed mice. We made extracellular recordings of single units from all layers of A1 using linear, multi-channel probes. A transgenic strategy was used to express channelrhodopsin 2 selectively in SST interneurons. Sounds were presented when SST neurons were either activated via laser or immediately following laser offset (i.e., deactivation). We characterized the effects of SST neuron activation and deactivation on regular- and fast-spiking A1 units through a linear classifier that allowed us to determine how reliably sound frequency was encoded over time. Optogenetic activation of SST neurons bidirectionally modulated cortical excitability, strongly suppressing activity during laser stimulation but producing rebound excitation upon deactivation. Neurons typically exhibited a short-latency, prolonged response to their best frequency (BF); however, spectrotemporal encoding of preferred stimuli was disrupted during both activation and deactivation of SST neurons. Deactivation of SST neurons unmasked subthreshold input that destabilized frequency tuning preference over time. Together, these changes degraded cortical decoding of sound frequency, particularly when the stimulus decoder was trained with longer integration windows. Finally, a sliding window decoder revealed that frequency processing during modulation of SST neurons diverges from normal such that the decoder fell to chance just 25 milliseconds after tone onset. These results suggest that cortical SST neurons are necessary to shape the temporal dynamics of cortical encoding of sound. SST neurons synapse onto dendrites and control both dendritic integration and active dendritic mechanisms. Disrupting their activity may occlude their ability to shape the intraneuronal processing required to generate a robust, time-insensitive representation of elementary sound properties.
Behavioral significance and auditory cortical representation of harmonic vocalizations

Nina So (1), Sarah M N Woolley (2)

Columbia University

Vocal communication relies on the brain’s ability to detect and process vocal sounds in the environment. Auditory systems that are optimized for vocalization processing must robustly encode acoustic features that distinguish vocalizations from other sound categories. One acoustic hallmark of speech and other animal vocalizations is harmonic structure. Sounds that are harmonic contain energy concentrated at integer multiples of a fundamental frequency, and higher harmonicity correlates with greater pitch salience. Our studies address two main questions. First, how does variation in harmonic structure alter the behavioral significance of vocal sounds? Second, where does selectivity for harmonic structure arise in the auditory processing pathway? We use zebra finches (Taeniopygia guttata) to address these questions. These songbirds use harmonic calls and songs for social communication and have well-defined, experimentally accessible auditory systems. Zebra finches exchange distance calls, a highly harmonic class of vocalizations, to maintain contact with each other when visually separated. Using a behavioral paradigm in which we present distance calls and assess birds’ vocal responses to these stimuli, we found that behavioral responsiveness is sensitive to the degree of harmonic structure in stimulus calls; birds respond less to spectrally degraded versions of calls compared to natural versions of the same calls. To test whether the responses of auditory cortex (AC) neurons are similarly sensitive to harmonic structure, we recorded the responses of single AC neurons to natural and spectrally degraded distance calls, in awake, restrained zebra finches while they listen to the calls. We recorded single unit responses from: 1) the intermediate region (Field L2), which receives information from the thalamus; 2) the superficial regions (Field L1/CM), which receive input from the intermediate region; and 3) the deep region (Field L3), which receives input from both intermediate and superficial regions. Preliminary results suggest that neuronal responses in the intermediate and superficial regions do not differ with variations in calls’ harmonic structure. In contrast, deep region neurons respond more to calls with high harmonic structure than to spectrally degraded versions of those calls. Together, results suggest that harmonic structure contributes to the behavioral salience of vocal communication sounds, and that neural selectivity for this feature arises in the deep region of primary AC.
Auditory Cortical Alterations in Mouse Models of Autism Spectrum Disorder

Tanya Gandhi (1), Charles C. Lee (1)

(1) Department of Comparative Biomedical Sciences, School of Veterinary Medicine, Louisiana State University, Baton Rouge, LA

Autism spectrum disorder (ASD) is a group of neurodevelopmental disorders characterized by deficits in social interaction and communication along with restricted interests and repetitive behaviors. Sensory processing abnormalities such as auditory perceptual impairments are frequently seen in the autistic individuals and might contribute to the complex behavioral symptoms associated with the disorder. In this study, we examine two genetic mouse models of ASD. The first model harbored homozygous null mutations in the contactin associated protein-like 2 gene (CNTNAP2-/-), associated with syndromic form of ASD known as the cortical dysplasia-focal epilepsy syndrome (CDFE). The second mouse model used recapitulates a deletion in human chromosome 16p11.2, implicated in 1% of individuals with ASDs (Brunner et al., 2015).

Using these genetic mouse models, we investigated the neuroanatomical alterations in the auditory cortex which might underlie cortical dysfunction during early postnatal stages into adulthood. One of the important neuropathological feature in these genetic mouse models of ASDs is the abnormal migration of neurons destined for superficial cortical layers to layer 5 of the cortex (Penagarikano et al., 2011), which could result in abnormal activity, thereby interfering with the normal sensory processing in the auditory cortex of these mutant mice. In our study, we employed labeling of auditory cortical neurons with CUX-1, which is a marker for neurons normally localized to the superficial cortical layers (II-IV). Additionally, we also immunostained the auditory cortex for the GABAergic cells marker GAD67 and one of the GABAergic cell type, parvalbumin (PV) interneurons. Deficits in auditory processing originating from synaptic dysregulation might also be due to the abnormalities in the extracellular matrix structures, such as the perineuronal nets (PNNs), analyzed by lectin staining in the auditory cortex of the mutant and the control animals. By quantifying the laminar distribution of CUX-1 positive cells in auditory cortex of these genetic models of ASDs, our findings show increase number of CUX-1 positive cells mislocalized to the lower layers (V and VI) of the auditory cortex as compared to the control animals. Furthermore, our results indicate a decrease in the PNN/PNN co-localized PVs positive cell density in the genetic mouse models in contrast to the controls. Also, there was no significant change in the number and laminar distribution of the cortical PV interneurons in the mutant and the control animals. These results suggest that the neuroanatomical alterations in the auditory cortex could play a role in the auditory processing impairments that account for some of the core autistic behaviors such as the hyper- and hypo- sensitivity to acoustic stimuli.
Response Types of the Superior Olivary Nucleus of the Barn Owl

Anna Kraemer (1), Catherine Carr (2)

(1) Neuroscience and Cognitive Sciences, University of Maryland, College Park
(2) Biology Department, University of Maryland, College Park

Inhibition is important for auditory processing. In chickens, the superior olivary nucleus (SO) provides inhibitory feedback to first order nuclei, and has been hypothesized to improve ITD coding (Fukui et al., 2010; Burger et al., 2011; Tabor et al., 2012). Few studies have characterized SO response types in barn owls. Moiseff & Konishi (1983) recorded from the owl SO in vivo, and found the majority of units were only excited by the ipsilateral ear. All units recorded from the SO were ILD insensitive, although the authors noted their search criteria might have increased the percentage of monaural response types (Moiseff & Konishi 1983). The SO is morphologically heterogeneous, however, and projects to multiple targets. Some SO neurons send inhibitory projections to two or more ipsilateral brainstem nuclei, while projections to the contralateral SON may originate from a separate population of neurons (Burger et al., 2005). Other SO neurons project to the inferior colliculus (Takahashi et al., 1988). Given the heterogeneity of SO cell types and projections, it seems likely to contain multiple response types. In order to examine auditory responses in SO, and to study the effects of descending inhibition provided to the first order auditory nuclei, we characterized barn owl SO responses in vivo. Electrolytic lesions confirmed recording locations. We measured tuning curves, tested for phase locking, and analyzed peristimulus time histograms, latency, regularity and PSTHs in addition to measuring rate-level function responses to noise stimuli. We also analyzed binaural measures of sensitivity to interaural level difference and interaural time difference, to categorize each response type. SO response types analyzed include primary-like, chopper, and onset, which are also found in the nucleus angularis (Koppl & Carr 2003). A majority of recorded SO single units were broadly tuned off-responses, similar to mammalian SPON units (Kulesza et al., 2003) and more responsive to noise than tones. We conclude that the SO displays similar response types and heterogeneity in response types to nucleus angularis, but more SO units preferred broad band stimuli, and many were binaural. These broadly tuned responses may serve to regulate the overall firing rate of other auditory brainstem nuclei and/or the contralateral SON.

Grant# NIDCD 000436
A novel mutual information estimator to measure spike train correlations in a model of the thalamocortical network

Ekaterina D. Gribkova (1), Baher A. Ibrahim (2), Daniel A. Llano (1,2)

(1) Neuroscience Program, University of Illinois at Urbana-Champaign
(2) Molecular and Integrative Physiology, University of Illinois at Urbana-Champaign

The impact of the various states of the thalamus on information transmission to the cortex remains poorly understood. This limitation exists due to the rich dynamics displayed by thalamocortical networks and because of the inadequate tools available to investigators to characterize those dynamics. Herein, we introduce a novel estimator of mutual information between spike trains and use this estimator to determine the impact of a computational model of thalamic state on information transmission. A limitation of previously employed mutual information estimators is their reliance on fixed partitions of spike trains. Therefore, we developed a mutual information estimator with an adaptive partition for signals with different time scales and used this estimator to compare input spike trains to output spike trains. It was observed that this estimator was superior to other mutual information estimators with fixed partitions when used to analyze simulated spike train data with different mean spike rates, as well as electrophysiological data from simultaneously recorded neurons, since it provided estimates that were closer to expected values and trends. When the estimator was applied to a thalamocortical model, it was found that thalamocortical (TC) cell calcium T-current conductance (T-current) influences mutual information between the input and output from this network. In particular, a T-current of about 50 nS appears to produce maximal mutual information between the input to this network (conceptualized as an afferent input to the TC cell) and the output of the network at the level of a cortical layer 4 neuron. Furthermore, at particular combinations of inputs to TC and thalamic reticular nucleus (TRN) cells, thalamic cell bursting correlated strongly with recovery of mutual information between thalamic afferents and L4. These studies suggest that a novel mutual information estimator using adaptive partitions has advantages over previous estimators, and that TRN activity can enhance mutual information between thalamic afferents and thalamorecipient cells in the cortex, while the cortex can recover mutual information between these thalamic afferents and thalamorecipient cells that is otherwise lost due to thalamic bursting.
Effect of focal attention on the cortical entrainment of the speech envelope

D. LESENFANTS (1), J. VANTHORNHOUT (1), E. VERSCHUEREN (1), L. DECRUY (1), B. SOMERS (1), T. FRANCART (1)

(1) ExpORL, Dept. of Neurosciences, KU Leuven, Leuven, Belgium

Recent progress in decoding brain responses to a speech stimulus has allowed developing novel measures for speech intelligibility based solely on electroencephalography (EEG). However, the cortical entrainment to speech varies over time independently of the intelligibility, reducing its applicability as an objective measure of speech intelligibility. We here hypothesized that periods with low-attention to the speech signal are associated with a lower cortical entrainment to speech, while higher levels of attention produces higher cortical entrainment. To evaluate this hypothesis, we presented running speech from a narrated story binaurally to fifteen normal hearing subjects. The reconstructed speech envelope was decoded from low-frequency EEG signals (0.5-4Hz? 64 electrodes) using a linear regression model (integration window: 0-250 ms) and the cortical entrainment was then computed as the correlation between presented and reconstructed speech envelope. In parallel, brain markers of focal attention were also extracted for each recording electrode using normalized spectral entropy. Finally, Spearman’s correlation between each channel’s spectral entropy and cortical entrainment over time was computed in the frontal area. Average cortical entrainment was 0.21 ± 0.08. Interestingly two participants presented lower cortical entrainment (respectively, 0.15 ± 0.06 and 0.14 ± 0.09) than the others (range: 0.20-0.27). These participants presented the highest correlation (respectively, 0.28 and 0.30) between the frontal entropy and the cortical entrainment (average: 0.03 ± 0.17) over time, suggesting a modulation of cortical entrainment with attention. Selection of periods with highest entropy (threshold was set as the 75th percentile) allowed to increase cortical entrainment to 0.18 (increase: 0.03) and 0.20 (increase: 0.06) respectively. Other individuals showed no difference in average in the cortical entrainment using only high frontal entropy periods. This study suggests that an increase in focal attention could induce an increase in the cortical entrainment to speech independently of the level of speech intelligibility. We here proposed to track and correct the effect of attention using spectral entropy-based markers in the frontal area. This could provide objective measures of speech intelligibility in clinical routine and open doors to automatic cochlear implant parameter adaptation without the intervention of a clinical expert.
Human speech is a sensory-motor process involving auditory self-monitoring to control vocal production and ensure accurate communication. Monitoring auditory feedback during vocal production allows one to quickly adjust speech to compensate for perceived changes in vocal output. Many animal species share similar feedback-dependent vocal control, however, the underlying neural mechanisms are unknown. Previous studies have demonstrated both a suppression of auditory cortex neurons during vocal production, and a sensitivity of such neurons to externally-perturbed auditory feedback. The role of this neural activity in vocal control is not known. We therefore investigated the responses of auditory cortical neurons in marmoset monkeys during vocal self-monitoring and feedback-dependent vocal control. Recent work has shown that, when presented with real-time frequency-shifted auditory feedback, marmosets will compensate for altered feedback by rapidly changing the acoustics of their ongoing vocalizations. We used implanted electrode arrays to record neural activities during self-initiated vocal production while simultaneously altering vocal feedback and observing the resulting vocal compensation. We found that neural activities in auditory cortical neurons were sensitive to altered feedback during vocalization, consistent with previous results. We further found that auditory cortical responses predicted subsequent changes in vocal production. These results demonstrate the function of auditory cortex in self-monitoring during vocal production, and suggest a role for the auditory cortex in feedback-coding and vocal motor control.
Temporal and rate coding of sound envelope and temporal fine structures of vocalizations in the primary auditory cortex of marmoset monkeys

Taku Banno (1,2), Wataru Suzuki (2), Naohisa Miyakawa (2,3), Toshiki Tani (4), Noritaka Ichinohe (2,4)

(1) University of Pennsylvania
(2) National Center of Neurology and Psychiatry
(3) National Institutes for Quantum and Radiological Science and Technology
(4) RIKEN BSI

Human speech and animal vocalizations are characterized by slow modulations in the sound envelope (ENV) and rapid fluctuations in temporal fine structures (TFS). Previous psychophysical studies demonstrated the ENV and TFS convey information of different aspect of speech; the ENV cues are critical for speech intelligibility, whereas the TFS cues are important for segregating the speech sound from the background noise. Although both of the acoustic cues are crucial for vocal communications, it has not been fully elucidated how the auditory system process this information.

To address this issue, we used marmosets as a model for human speech perception and conducted electrophysiological recordings from the primary auditory cortex (A1) of anesthetized animals while presenting acoustic stimuli consisting of various combinations of amplitude (AM) and frequency modulations (FM). The fundamental frequency and the depth of the modulations were matched to the marmoset vocalizations so that the stimuli sounded similar to their calls when the AM and FM were equated to the temporal parameters of the vocalizations.

We found the spike timing of the A1 neurons precisely followed the modulation frequency of the stimuli if the AM and/or FM were slower than ~10 Hz, suggesting the neurons temporally encode the slow rhythms in both the ENV and TFS of the sounds. A greater number of spikes were observed in the sounds with the slow AM and FM combinations, but the firing rate showed more obvious tunings for FM than AM as a population. The information carried by the firing rate was further examined by mutual information analysis, and we confirmed the general tendency that the A1 neurons encoded a higher amount of FM information. Interestingly, the neurons could have a high amount of information either for AM or FM but not for both, implying the A1 neurons were optimized for encoding the animal vocalization sounds. These results suggest the A1 neurons process information about ENV and TFS cues differently by the temporal and rate coding in a way that they efficiently encode the temporal structures in the conspecific vocalizations.
Missing Fundamental Pitch Perception with Semitone Precision in Marmosets

Xindong Song (1), Michael S Osmanski (1), Xiaoqin Wang (1)

(1) Johns Hopkins University

The perception of the missing fundamental, in which one perceives the fundamental frequency (F0) of a harmonic complex tone even when the F0 is removed, is one of the most important properties of pitch perception in humans. Some non-human species have also been shown to be sensitive to the pitch of missing fundamental sounds, although these previous animal studies have not demonstrated that this sensitivity to pitch has a precision comparable to that of humans. For humans to perceive Western musical melodies, a precision of at least one semitone is necessary. Marmosets (Callithrix jacchus), a highly vocal New World monkey species, have been shown to be able to discriminate a pitch difference of one semitone for harmonic complex tones with a periodicity of 440Hz (the standard A or A4 in ISO16). We hypothesized that marmosets also possess missing fundamental pitch perception with one semitone precision for the periodicity at 440Hz (A4).

In the current study, marmosets were trained to discriminate harmonic complex sounds that differed in both periodicity (F0) and the presence of an F0 component. The animals exhibited a high hit rate when a probe sound differed from a reference sound only in periodicity. In contrast, the animal’s hit rate was indistinguishable from the false alarm rate when a probe sound differed from the reference sound only in its F0 presence. These results show that marmosets discriminated these harmonic complex sounds based on their periodicity but not the presence of an F0 component, which suggest that they can perceive the missing fundamental pitch. For a periodicity of 440Hz (A4), marmosets could perform this task even when the pitch change was as small as one semitone. Whereas for a periodicity of 220Hz (A3), the pitch change had to be increased to two semitones. Such an increase is also consistent with their perceptual sensitivity to temporal envelope cue changes on A3, but not A4. This difference in sensitivity suggests that the mechanisms for missing fundamental pitch perception change between A3 and A4. Together, our findings demonstrate that marmosets have the capacity to process missing fundamental pitch similar to humans with a precision of at least one semitone for the periodicity above A4. This is the first time that a non-human species has been shown to have the ability to discriminate the missing fundamental pitch at this precision, which suggests that marmosets may potentially be able to discriminate musical melodies.
A reexamination of click detection during statistical learning

Tobias Overath (1)
David LK Murphy (1)

(1) Duke University

Click-detection during segmentation of continuously presented nonwords has been proposed as an online measure of statistical learning during language acquisition. In such a task, subjects are instructed to respond to clicks superimposed over a stream of tri-syllabic nonwords. In a previous study, reaction times (RT) to clicks paired with syllables within words (syllable 2, S2) slowed more rapidly than clicks paired with syllables between words (syllable 1, S1; Gomez et al., 2011), over a 4 min exposure period. Later work (Franco et al., 2015) demonstrated that it is unclear if this effect is a reflection of statistical learning of word boundaries or of competing attentional demands during active or passive word segmentation. It is also unclear if post-exposure measurements of statistical learning directly represent the efficiency of the segmentation process, subjects’ short-term auditory memory, or a combination of both.

Participants in this study listened to a stream of concatenated tri-syllabic non-words and were asked to detect clicks that were either placed between words (S1) or within words before the third syllable (S3), for which the composite within-word predictability is highest (previous studies placed within-word clicks before S2). The exposure time was also increased to 6 minutes to test the stability of statistical learning. The shift in click position to S3 was motivated by additional exploration of target detection during word segmentation: under explicit and implicit learning conditions, during which subjects were instructed to actively segment word boundaries, RTs were fastest during detection of target syllables occurring as the last in a trisyllabic word (Batterink et al., 2015).

Here, we show that detection of S3 clicks is faster than S1 clicks during explicit learning, replicating the behavior observed in syllable detection tasks (Batterink, 2015), but contradicting previous studies comparing clicks at positions S1 and S2 (Gomez et al., 2011; Franco et al., 2015). This suggests that one syllable is insufficient to trigger learning-guided shifts in attentional demand, but two syllables, when part of a statistically generated model, have an immediate salience effect. We also find that in the 2AFC word recognition stage, accuracy is correlated with the relative changes in S1 and S3 RTs, and that differences in S1 vs. S3 RTs diminished during minutes 5-6.

References:
High gamma neural responses dissociate between the acoustic and linguistic analysis of temporal speech structure


(1) Department of Neurosurgery, Duke University Medical Center
(2) Duke Institute for Brain Science, Duke University
(3) Department of Neurology, Duke University Medical Center

Speech perception entails the mapping of the acoustic waveform to its linguistic representation. For this mapping to succeed, the speech signal needs to be tracked across a large temporal range at high temporal precision in order to decode linguistic units (e.g. phonemes, syllables, words). Here we test how cortical processing of such temporal speech structure is modulated by higher-order linguistic analysis. To control the temporal scale of analysis, we used a novel sound-quilting algorithm that controls acoustic structure at different temporal scales; using fMRI, we recently showed that activity in human superior temporal sulcus (STS) increases as a function of temporal scale in an unfamiliar language (Overath et al., 2015). To control the linguistic content, we constructed speech quilts from both familiar and foreign languages. This ensures that any changes at the signal-acoustics level affect both languages identically, while manipulating the linguistic percept differently. Thus, neural responses that vary as a function of segment length but are shared or similar across the two languages suggest analysis at the signal-acoustics level, whereas neural responses that differ based on language familiarity imply the presence of linguistic processing.

Here, we recorded electrocorticography (ECoG) from electrodes placed over left temporal or fronto-temporal lobes in three patients who were undergoing pre-surgical monitoring for pharmacologically resistant epilepsy. Patients listened to 6 s long English or Korean speech, quilted with 30 ms or 960 ms segment lengths.

Electrodes with significant auditory responses were initially assessed via a permutation test between the 1 s time window following sound onset compared to the pre-stimulus baseline. Neural signals were filtered between 70 and 150 Hz (high gamma) and the results were Bonferroni corrected for multiple comparisons across electrodes. 20/124 electrodes demonstrated a significant auditory response. Within these electrodes, a follow-up analysis showed that sustained high gamma responses throughout the 6 s sounds showed a main effect of segment length (30 vs. 960 ms) in 85% (17/20), a main effect of Language (English, Korean) in 60% (12/20), and an interaction in 50% of auditory electrodes (10/20). Specifically, electrodes that showed an interaction generally displayed a larger increase in high gamma power as a function of segment length in English than in Korean. These results suggest that high gamma neural responses are a potential neural mechanism for tracking speech-specific temporal structure.
From riding a bike to learning a language, children are considered to be better skill learners than adults. One possible explanation for this superiority is that sensory systems display heightened plasticity during development, and thus may be more responsive to training. Developmental studies have documented differences between juvenile and adult learning, and have identified critical periods during which neural properties are particularly sensitive to environmental experience. However, the neural plasticity that accompanies juvenile learning remains unexplored. To address this issue, we recorded telemetrically from left auditory cortex (ACx) of juvenile gerbils as they trained and improved on an amplitude modulation detection task. Neurometric and psychometric sensitivity were simultaneously tracked across days. As performance improved, juvenile ACx units displayed significant within-animal correlations between neural and behavioral thresholds. These data will be compared to similar recordings from mature adult animals to evaluate whether or not plasticity mechanisms differ.
Neural Responses to Behaviorally Relevant Syllable Sequences in the Ferret Auditory- and Frontal Cortices

D. Duque(1), N. Joshi(1), D. Elgueda(1), J. B. Fritz(1) and S. A. Shamma(1,2)

(1) Neural Systems Lab, Institute for Systems Research, University of Maryland, College Park, MD, USA.

(2) Équipe Audition, Laboratoire des Systèmes Perceptifs, École Normale Supérieure, Paris, France.

Speech unfolds in time, and words and sentences are built around sequential strings of phonemes. As a result, spoken language relies not only on the ability to segregate and recognize individual phonemic elements in ongoing speech, but also in understanding that ordered combinations of such discrete elements form auditory objects. The generation of these auditory objects, known as words, is essential for language since once we create a sound-to-meaning association, oral communication relies on recalling these ‘word memories’ when a complex sound input is processed. Such memories may be stored in association cortices and retrieved when a sound input matches an acoustic memory template generated in auditory cortex.

To explore the neural mechanisms by which the auditory system encodes human words, we train ferrets to discriminate tri-syllabic consonant-vowel words. Ferrets are an excellent animal model to study speech representation because their hearing is similar to humans, their auditory cortex is complex enough to encode all phoneme classes (Mesgarani et al., 2008) and they can be trained to differentiate and recognize syllable sequences (Bizley et al., 2015; Duque et al., 2016). Several ferrets were trained on a conditioned avoidance GO/NO-GO task to lick a spout for water upon presentation of distractor words (e.g. FA-BE-ti) but to refrain from licking the spout after a target word (e.g. FA-BE-LO) was presented, in order to avoid a mild shock. Once the animals learned the task, a head-post was implanted to enable single unit electrophysiological recordings in primary auditory cortex (A1), higher order auditory cortical areas (dPEG) and frontal cortex (FC) while the animals were performing the behavioral task.

Preliminary neurophysiological data indicate that, unlike the stimulus-dependent responses in A1, responses in dPEG reveal differences between target and distractor words based on associated behavioral meaning. This study suggests that neurons in higher order auditory cortical areas can encode specific syllable sequences. These results will help us understand the neural basis for the representation of complex sound sequences, and yield deeper insight into neural mechanisms underlying initial stages in human speech processing.
Attentional effects on neural responses in a cocktail party model in the ferret auditory cortex

N. Joshi(1,2), D. Duque(1), J. B. Fritz(1) and S. A. Shamma(1,2,3)

(1) Neural Systems Lab, Institute for Systems Research, University of Maryland, College Park, MD, USA.

(2) Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA

(3) Équipe Audition, Laboratoire des Systèmes Perceptifs, École Normale Supérieure, Paris, France.

During the cascade of neural responses triggered by sound from multiple sources hitting the tympanum, the brain needs to deconstruct the mixture by auditory scene analysis into different streams and then select the most relevant streams for further processing. Many different ideas have been proposed as to how the brain segregates sound and decodes the auditory scene, some suggesting that scene segregation is pre-attentive, activating separate neural populations that respond to different auditory attributes, while others emphasizing the importance of attention. In this work, we study the effect of attention on neural representation in a multi-speaker scenario in the ferret auditory cortex.

Ferrets are an excellent animal model to study speech representation because the range of sounds frequencies heard by them is similar to humans, their Auditory Cortex is complex enough to encode phoneme classes (Mesgarani et al., 2008) and they can be trained to differentiate syllable sequences (Bizley et al., 2015; Duque et al., 2016). To explore the role of attention in neural mechanisms of stream segregation, we trained ferrets to discriminate tri-syllabic pseudo-words using a conditioned avoidance GO/NO-GO task. We also trained them to attend to a target word (Eg. FA-BE-KU) by a female speaker while ignoring a simultaneous male background speaker. In order to analyze neural encoding in stream segregation, we perform single and multi-unit neurophysiological recordings in primary and secondary areas of the Auditory Cortex. We explored whether responses during the tasks showed adaptively enhanced representation of the attended female speaker.

Preliminary neurophysiological data indicate that neuronal responses to words in the male distractor voice in the cocktail party scenario showed specific suppression, enhancing relative representation of the attended female target word. This study shows that neurons in higher order Auditory Cortex can encode pseudo-words selectively in a multi-speaker scenario. These results will help us understand the neural basis for representation of complex sound sequences and yield deeper insight into neural mechanisms underlying initial stages in human speech processing of the cocktail party problem.
Abnormalities in cortical parvalbumin-positive interneuron density and distribution and auditory evoked potentials in a mouse model of 22q11.2 Deletion Syndrome

Fhatarah A Zinnamon (1,2), Freya G Harrison (2), Kuan Hong Wang (1), Jennifer F Linden (2)

(1) Unit on Neural Circuits and Adaptive Behaviors, National Institute of Mental Health, NIH

(2) Ear Institute, University College London

22q11.2 Deletion Syndrome (22q11DS) is the strongest known genetic risk factor for the development of schizophrenia (SCZ). The Df1/+ mouse model of 22q11DS recapitulates many features of human 22q11DS and schizophrenia, including cognitive impairment and frequent otitis media, a middle ear disease that can cause conductive hearing loss. Both hearing loss and SCZ risk factors have been associated with abnormalities in parvalbumin-positive (PV+) inhibitory interneuron circuitry in the cortex. Additionally, impairments in auditory evoked potentials (AEPs) serve as endophenotypic markers of SCZ. However, the relationship between hearing loss, genetic risk of SCZ, AEPs, and PV+ interneuron circuitry remains poorly understood. We explored this relationship through immunohistochemical and electrophysiological studies of auditory and frontal cortices in wildtype (WT) mice and Df1/+ mice with and without hearing loss. We tested hearing thresholds using auditory brainstem response measurements. During the same session, we also recorded AEPs to explore how hearing loss and genotype affect auditory cortical activity. PV+ immunohistochemistry on coronal sections through the auditory and frontal cortices of the mice indicated significant reductions in PV+ interneuron cell counts and densities in Df1/+ mice within the primary auditory cortex (A1) but not the secondary motor cortex (M2). Quantifications of PV+ interneuron density across cortical layers showed that PV+ cell distributions were abnormal in both A1 and M2. Additionally, Df1/+ mice with hearing loss displayed altered AEPs suggestive of increased central auditory gain compensating for reduced input. Results indicate that genetic risk of schizophrenia and developmental hearing loss interact to produce cumulative abnormalities in PV+ interneuron networks with functional physiological sequelae.
How many sound locations can humans distinguish at a time? Implications for neural processing of auditory space

L. Farrell (1), J. M. Groh (1)

(1) Duke University

How the brain encodes multiple sounds is unclear, particularly given recent evidence that sound azimuth is encoded in the level of activity in a neural ensemble rather than via an auditory space map of circumscribed receptive fields. One possibility is that the brain takes advantage of frequency differences to sort different sound sources to different regions of tonotopic maps (Willett et al. Soc Neuro Abstr 2016). Alternatively, the brain may use time division multiplexing to encode multiple sounds via fluctuating activity rates corresponding to individual sounds across time (Caruso et al., biorxiv, 2017).

Insight into which possibility or possibilities may apply and how they operate can be gained from knowing how many sounds humans are capable of distinguishing at one time. In this study, 8 different bandpass filtered noise stimuli (with different center frequencies) were combined and presented simultaneously from between 1 and 8 different locations. For example, all 8 bandpass filtered stimuli were combined and played from 1 speaker, all were separately played from 8 speakers, or any combination in between. Human participants were asked to report how many different locations they heard on each trial. Since the frequency information was held constant, we interpret these results chiefly under the time division multiplexing theory.

We found that the reported number of locations scaled with actual number up to about 4 sounds. However, participants tended to underestimate the actual number and were accurate only for 1-2 locations. We conclude that the time division multiplexing component of multiple stimulus encoding may operate on a duty cycle that permits representation of a maximum of four sound locations at a time.
Decoding the cortical tracking of auditory motion using EEG

Adam Bednar (1),(2) and Edmund C. Lalor (1),(2)

(1) School of Engineering, Trinity Centre for Bioengineering and Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin 2, Ireland.

(2) Department of Biomedical Engineering and Department of Neuroscience, University of Rochester, Rochester, NY 14627, USA.

It is of increasing practical interest to be able to decode the spatial characteristics of an auditory scene from electrophysiological signals. This is particularly relevant for cognitively controlled hearing aids and for the evaluation of virtual acoustic environments. However, the cortical representation of auditory space is not well characterized -- particularly as it relates to auditory motion -- and whether it is possible to decode cortical signals using non-invasive neuroimaging techniques remains unclear. In previous work (Bednar et al., 2017) we showed that we can classify evoked electroencephalographic (EEG) responses to discrete static sound stimuli as a function of their location in space. In the present study, we extend this work with the aim of decoding the trajectory of continuous, moving auditory stimuli. Subjects listened to pink noise over headphones, which had been spectro-temporally modified to be perceived as randomly moving on a semi-circular trajectory in the horizontal plane. Participants were asked to respond to infrequent tremolo targets which were embedded in the stimuli. While subjects listened to the stimuli, we recorded their EEG using a 128-channel acquisition system. The data were analyzed by 1) deriving a linear mapping, known as a temporal response function (TRF), between the stimulus and a training set of EEG data, and 2) using this TRF to reconstruct an estimate of the time-varying sound source azimuth from a test set of EEG data. Results show that we can decode sound trajectory with a reconstruction accuracy significantly above chance level. Specifically, we find that both the phase of low frequency (<4Hz) cortical activity and the power of higher frequencies (8-12 Hz) can predict the sound source azimuth. Further analysis of the reconstruction model parameters reveals strong lateralization at different scalp regions for different EEG frequencies.
Categorical memory representation in ferret auditory and frontal cortices

Pingbo Yin1(1), Jonathan B. Fritz(1-2), Shihab A. Shamma(1-3)

(1) Neural Systems Laboratory, Institute for Systems Research, (2) Electrical and Computer Engineering Department, University of Maryland, College Park, Maryland, USA. (3) Sensory Perception Laboratory, Ecole Normale Superieure, Paris, France

Categorization can arise by grouping of sensory stimuli sharing perceptual similarity or by training in which sensory stimuli are assigned to behaviorally relevant classes associated with learned responses. Neuronal category representations have been previously identified in vision in prefrontal, inferotemporal and parietal cortices, but a clear demonstration in audition has been elusive. In previous work, we showed that ferrets could be trained to classify auditory stimuli distributed along continuous feature dimensions (e.g. frequency or amplitude modulation) into classes based on behavioral meaning (Go or No-Go). Here we explore how these categories become represented in different fields of ferret auditory cortex (A1 and secondary areas in dPEG) and frontal cortex (FC), when animals were passively listening to sounds or actively engaged in task. To investigate the underlying mechanisms of category representation, we distinguish between ‘category-selective coding’ and ‘sensory-selective coding’. Specifically, we expect stimuli from the same category to evoke shared response patterns across the neural population despite potentially large sensory differences between them, e.g., in perceived pitch or modulation. Two metrics based on the Euclidean distance of the population response between stimulus pairs were computed to evaluate these categorical effects: (a) categorical index (CI) - defined as the difference between the mean distances of the pairs the stimuli from different categories and of the pairs the stimuli from same category; and (b) the proportion of variance explained by the categories. This analysis revealed that categorical information (Go versus No-Go) was observed in all three regions during task performance. Furthermore, the dynamics of this representation reveals that the categorical information emerges earliest in FC neurons, followed in dPEG, and then in primary A1. These results are consistent with the findings from visual system and provide insight into how the primary and secondary auditory cortices are differentially involved in classifying acoustic inputs during active auditory memory retrieval.
Effects of voice continuity and stimulus rate on auditory working memory

Sung-Joo Lim (1), Barbara G. Shinn-Cunningham (1), Tyler K. Perrachione (1)

(1) Boston University

Speech perception can be cognitively demanding when incoming speech varies from token to token. The auditory system adapts to speech characteristics, leading to more rapid and accurate perception for speech from a single talker compared to varying talkers. To date, facilitatory effects of speaker adaptation have used tasks requiring immediate recognition of speech. As a result, it is not clear whether speaker adaptation improves encoding or recall of speech information being maintained in auditory working memory, or how effects change when listeners have more time to process each individual speech token. Using a delayed recall of digit span task, we investigated whether time between speech tokens mitigates interference of processing multiple talkers, e.g., by allowing for memory rehearsal of a token before a subsequent token has to be processed and encoded. On each trial, listeners encoded a seven-digit sequence and recalled it after a 5-s delay. Sequences were spoken by either a single or multiple different talkers at three different presentation rates (0, 200, or 500 ms inter-digit intervals). Recall was less accurate and slower for sequences spoken by multiple talkers than by a single talker. The difference between processing single vs. multiple talkers decreased as the presentation rate decreased. However, slower rates of sequence encoding generally improved listeners’ memory recall overall. These results suggest that the speaker adaptation effect persists in auditory working memory items, especially when listeners are not provided with time to encode individual tokens before a subsequent token is heard. These behavioral results provide the foundation for further investigations of the electrophysiological correlates of how speech variability influences auditory encoding and working memory processes.
Integrating behavioral context into auditory encoding models

Luke A. Shaheen (1), Sean J. Slee (2), Stephen V. David (1)

(1) Oregon Health & Science University
(2) Biotronik

The spectro-temporal receptive field (STRF) and more sophisticated variants have emerged as a standard means of characterizing representation throughout the auditory system. One appealing quality of encoding models is their generalizability—that they can be evaluated and compared according to how well they predict the time-varying neural response to any arbitrary sound. Most existing models describe neurons as static sensory processors. However, behavioral context, reflected in the acoustic environment or behavioral state of the listener, can influence sound-evoked activity, particularly in central auditory areas.

We developed a computational framework for unbiased comparison of a large number of linear and nonlinear encoding models derived from the STRF. We identified models with minimal complexity, i.e., those which required as few free parameters as possible but were still able to optimally predict sound-evoked responses. We then used models with optimal complexity to analyze single-unit neural data collected from inferior colliculus (IC) and primary auditory cortex (A1) of awake ferrets while they switched between passive listening and a tone detection task. We fit models using stepwise regression, in which individual free parameters were either fixed across behavioral states or allowed to vary between states. In A1, behavior-dependent changes in encoding occurred in spectral-temporal selectivity as well as overall excitability of neurons. In IC, the majority of changes were accounted for mostly by changes in excitability. The latter model required many few behavior-dependent parameters. The increasing complexity of context effects between IC and A1 supports a hierarchical model in which task-related plasticity emerges in IC and is inherited by A1. Ongoing work is exploring the simultaneous integration of multiple behavioral state variables on neural coding (task engagement, arousal, attention) and effects of direct circuit manipulations on context-dependent encoding properties.
Cross-modal interactions in thalamo-thalamic reticular interconnectivity

Kush Paul (1), Daniel A. Llano (2)

(1) University of Illinois at Urbana-Champaign
(2) University of Illinois at Urbana-Champaign

The thalamic reticular nucleus (TRN) consists of GABAergic neurons that have been hypothesized to modulate the flow of sensory or motor information through first- and higher-order thalamic nuclei. The TRN has closed loop reciprocal connectivity with the thalamus which are well established, however, there is a growing evidence of non-reciprocal or open loop connectivity between them. While the majority of information transmission to and from the cortex occur via separate thalamic nuclei corresponding to the sensory modalities, there has also been evidence of crosstalk in the thalamus between the modalities which occur through an inhibitory disynaptic pathway involving the TRN. Using laser scanning photostimulation with glutamate uncaging we have investigated the patterns of reciprocal and non-reciprocal thalamus-TRN interconnectivity in the colliculo-thalamocortical and somatosensory slice preparations.

Whole-cell patch-clamp recordings were done in cells in discrete thalamic nuclei which were obtained in the colliculo-thalamocortical and somatosensory slice orientation from mice (P12-24). Recordings were obtained using standard extracellular solution and Cs based intracellular solution in the voltage clamp mode. The holding potential was kept at 0 mV or +10 mV to maximize amplitudes of the photostimulation evoked inhibitory postsynaptic potentials. MNI-caged glutamate (Tocris) was added to recirculating ACSF and stimulated using a pulsed UV laser (355 nm, DPSS) over a grid of points (32 by 32 array) encompassing TRN and several thalamic nuclei (MGB, VB, dLGN) and focal photolysis was accomplished by non-neighbor stimulation of points within the grid. Current responses were obtained in the voltage clamp configuration with the outward current amplitude representing strength of the disynaptic connection.

We found strong disynaptic connectivity between neurons in the thalamus and TRN which were topographically organized. We also found disynaptic connectivity between nuclei in the thalamus which suggest the presence of open loop connections and cross-modal interconnectivity with the thalamus.
Task-related plasticity in the ferret IC during tone detection in noise

Daniela Saderi (1), Stephen V. David (1)

(1) Oregon Health & Science University

The inferior colliculus (IC) is a midbrain region that integrates ascending auditory information and receives top-down feedback from thalamus and cortex. Previous work in our laboratory found that engaging in a simple tone-detection task modulates neural responses in the ferret IC, and this modulation was found to be predominantly a suppression of evoked responses during behavior compared to passive listening. Auditory behavior involves multiple aspects of internal state, and the current study sought to isolate effects of task difficulty on sound representation in the IC. We trained ferrets on a tone-in-noise detection behavior, where animals had to detect a tone embedded in the noise (target) and ignore the noise alone (distractors). Changing the signal-to-noise ratio of the tone relative to the noise (target SNR) controlled task difficulty. Trial blocks alternated between ‘difficult’ (low target SNR) and ‘easy’ (high target SNR). The task also included occasional probe targets, presented at fixed SNR. Probes allowed us to isolate the effects of changing expected SNR (difficult or easy block) on behavioral performance as well as on sound-evoked neuronal activity. We hypothesized that behavioral performance in difficult blocks measured at the probe stimuli would be higher than in easy block as a result of increasing effort. However, we observed no consistent difference in perceptual sensitivity or criterion between conditions, suggesting that the ferrets do not alter their behavior across SNR conditions. We performed single-unit tetrode recordings in the IC during easy and difficult behavioral conditions as well as during passive listening to task stimuli. Our preliminary results did not reveal any modulation of sound responses between behavioral conditions. However, in contrast to the suppression previously observed during pure tone detection, we found that neural responses to both distractor and target stimuli were enhanced during tone-in-noise behavior. To explore this discrepancy further, we have introduced a pure tone condition to our behavior, where the animal has to detect a tone alone. Thus, we will be able to measure if detection of tone alone versus tone embedded in noise differentially modulates activity of the same IC neuron or if these oppositional effects are rather the result of behavioral training on the two behaviors, potentially reflecting two different learning strategies.
Psychophysics of auditory streaming based on spectral and spatial cues in rhesus monkeys

Jaejin Lee (1), Taku Banno(1), Yon I. Fishman(2), Yale E. Cohen(1)

(1) University of Pennsylvania
(2) Albert Einstein College of Medicine

A fundamental aspect of hearing is parsing auditory stimuli into discrete perceptual representations of sounds in the environment. The auditory system accomplishes this parsing, in part, by grouping acoustic stimuli with similar spatial or and spectral features into one perceptual stream and separating those with different features into different streams. However, the degree to which spectral and spatial cues contribute to sound-stream segregation has not been fully elucidated, especially in non-human primates.

Here, we tested how the degree to which spectral and spatial cues affected the ability of rhesus monkeys to effectively group or segregate acoustic stimuli into separate auditory sound streams by conducting a series of psychophysical experiments. Specifically, we presented two sequences of alternating interleaved tone bursts that varied in frequency, sound level, and spatial separation from a free-field speaker array. Monkeys were trained to report hearing a deviantly loud tone burst. This report served as a proxy for stream segregation: by design, this deviant could only be detected if the two sequences of tone bursts were perceived as separate streams.

As expected, we found that monkeys’ ability to detect the deviant tone greatly improved as the frequency difference between the two tone-burst sequences increased. Likewise, as the spatial separation between the two sequences increased, the ability to detect the deviant tone was also substantially enhanced. Interestingly, spatial separation, in this paradigm, was a more potent segregation cue than spectral separation. Interestingly, spatial cues, in this paradigm, served as were a more potent segregation cue than spectral cues. Furthermore, spatial and spectral cues interacted synergistically to facilitate deviant detection and hence stream segregation. Psychophysical findings are consistent with human performance on a comparable task. This study provides the first demonstration of the combined effect of spectral and spatial cues on auditory streaming in non-human primates.
103 Ultrasonic Noise Elicits Significant BOLD Signal in the Auditory Brainstems of Echolocating Bats

Stuart Washington¹, Herbert Peremans², Jan Steckel³, Rukun Hinz¹, Tom Venneman¹, Lisbeth Van Ruijsevelt¹, Elisabeth Jonckers¹, Jan Scholliers⁴, Sayuan Liang¹, Johan Van Audekerke¹, Marleen Verhoye¹, Georgios Keliris¹, Karl-Heinz Esser⁵, and Annemie Van der Linden¹

1. Bio-Imaging Lab, University of Antwerp, BEL; 2. Department of Engineering Management, University of Antwerp, BEL; 3. Department of Electrical Engineering, University of Antwerp, BEL; 4. Department of Biology, University of Antwerp, BEL; 5. University of Veterinary Medicine, Hannover, GER.

Echolocating bats were the topic of classical neurophysiological studies that provided substantial knowledge about auditory processing within mammalian nervous systems. These highly social species orient by generating bio-sonar signals and communicate with conspecifics via acoustically complex vocalizations. Such deep reliance upon acoustic signals likely drove the evolution of hypertrophic neuroanatomical structures and systematic representations of stimulus features throughout their auditory systems. Single unit recording and other classical neurophysiological methods are limited in their collective ability to non-invasively detect either neural activity throughout or functional interactions between these hypertrophic auditory processing structures. On the other hand, functional MRI is ideally suited for non-invasive, global explorations of auditory processing at a gross neuroanatomical level. Here, we used a 7 Tesla MRI scanner to acquire echo-planar images [voxel size = (469 x 469 x 1100) µm³; TR/TE = 2000ms/18ms; field of view = (30) mm²; flip angle = 90°; matrix size =64 x 64] of blood oxygenation level dependent (BOLD) signal changes elicited by the presentation of ultrasonic white-noise (20-120 kHz) in 9 pale spear-nosed bats (Phyllostomus discolor). Specifically, our stimuli were presented through non-magnetic, solid-state ultrasonic speakers (TAKET™ Batpure) and consisted of 50 blocks (16 secs) of white noise ripples (on: 5-ms; off: 10-ms) and 50 blocks of silence. White noise largely overlapped the bat’s range of hearing (5-130 kHz), which only slightly overlapped with the highest frequency harmonics of MRI scanner noise (15 kHz). Anesthetic induction with isoflurane preceded neuroimaging; neuroimaging coincided with medetomidine infusion (bolus: 0.025 mg/kg; infusion: 0.05 mg/kg/hr). Relative to silence, ultrasonic noise elicited significant (cluster-level \[p_{FWE} = 0.009\]; cluster size \[K_E = 475\]) BOLD activity in the auditory brainstem, most notably the inferior colliculus and cochlear nucleus. Significant clusters (cluster-level \[p_{FWE} < 0.001\]; cluster size \[K_E = 969\]) emerged in the auditory cortex when comparing silence relative to noise. These data suggest that the same broadband ultrasonic noise that elicited activation within the auditory brainstem suppressed activation in the auditory cortex. A large sinus cavity caused considerable susceptibility artefact in the frontal lobe but spared other brain regions. These results suggest that auditory fMRI studies of echolocating bats have the potential to reveal systematic stimulus feature representations within
neuroanatomical regions of the auditory system and the physiological interactions between them.
104 Perceptual categorization of pup vocalizations in the auditory cortex of maternal mice

Jennifer K. Schiavo1,2,3, Robert C. Froemke1,2,3

1. Skirball Institute of Biomolecular Medicine, New York, New York. 2. Sackler Institute of Graduate Biomedical Sciences, New York, New York. 3. Departments of Otolaryngology and Neuroscience/Physiology, New York, New York.

Learning requires generalization from specific exemplars to enable animals to respond reliably to different stimuli with the same behavioral significance. Perceptual categorization of vocalizations facilitates communication by grouping stimuli into discrete categories despite variability in acoustic features across individuals. Little is known about the mechanisms underlying learned category formation at the single-cell and population level, as well as the link between categorical representations and generalization from learned categories. Here we take advantage of a natural behavior in adult mice: retrieval of nest-isolated pups based on distress calls (Ehret Nature 1987, Liu & Schreiner PLoS Biol 2007, Cohen et al. Neuron 2011, Marlin et al. Nature 2015). While mothers rapidly respond to distress calls by bringing the pup back to the nest, naïve virgins initially do not behaviorally respond to these calls. USV sequences are variable between and within individuals, however the behavioral significance is the same: to signal the mother. This behavior requires females to perceptually categorize USVs, and serves as a model to examine the neural mechanisms underlying category formation in the auditory cortex. Here, we used behavioral measurement and 2-photon imaging to monitor changes in neural and behavioral categories as virgin females gain maternal experience.

First, we assessed behavioral responses in maternal mice to natural and morphed USVs (in the temporal and pitch domains) using a Y-maze. Mothers reliably approached speakers playing pre-recorded USVs, as well USVs temporally morphed within the natural range (i.e., a rate emitted by pups). Using 2-photon imaging in awake mice, we assessed neural response boundaries for pup calls in maternal and non-maternal females. Excitatory neurons in maternal auditory cortex responded invariantly to time-warped USVs within the natural range, while excitatory neurons in naïve females did not respond invariantly to morphed calls. The behavioral and neural data thus far is indicative of distinct cortical differences in the processing of USVs between maternal and naïve females, providing an opportunity to further study how the brain shifts from a non-maternal to maternal state to enable parental behaviors.