

1. Recalibration of excitatory and inhibitory local cortical networks supports neural and perceptual recovery of simple – but not complex – sound processing following a long-lasting reduction in auditory input strength

Jenifer Resnik, Daniel B. Polley

The auditory system employs a variety of rapid gain control mechanisms to adjust neural coding sensitivity to match transient shifts in acoustic signal energies. In addition to these “fast acting” gain control systems, central auditory neurons also exhibit slower gain control that adjusts neural excitability following long-lasting reductions in auditory input strength, for example, a deprivation of afferent input from the ear. While there is a general notion that increased “neural amplification” following a partial blockade of input from the ear is enabled by changes in inhibitory strength, the time course and cell type specific circuitry modifications that underlie slow changes in central auditory gain remain unknown.

Here we performed chronic, cell type-specific 2-photon calcium imaging to simultaneously visualize sound-evoked GCaMP signals in genetically identified inhibitory PV (parvalbumin expressing) neurons alongside neighboring PPy (putative pyramidal) cells in the auditory cortex of awake adult mice, before and after a controlled loss of afferent input from the cochlea. This approach allowed us to track the daily dynamics in identified cell types, at different spatial scales – single cell to network activity, and temporal scales – hours to weeks following peripheral insult. We found an increase in spontaneous activity in PPy cells on the day of the insult followed by changes in spontaneous activity of PV cells. Both excitatory and inhibitory cells exhibited a major decrease in toned evoked responses, which recovered almost back to baseline levels two weeks post injury. For more temporally complex stimuli, such as tones embedded in background noise, both PPy and PV cells showed an increase in response thresholds that didn’t recover during the two weeks that followed the insult.

Our imaging data demonstrated complete cellular and network recovery for simple stimuli, but persistent coding deficits for more complex stimuli such as tones in noise. To explore the perceptual implications of these observations, auditory operant behavioral measurements were performed in head-fixed mice before and two weeks following damage to cochlear nerve afferents. As predicted from our imaging data, mice showed complete perceptual recovery for detecting tones in silence despite a massive loss of auditory nerve input. However, tone detection in noise remained impaired.

Collectively, our work provides new insight into slow compensatory plasticity in PV and PPy neurons in the auditory cortex that restores neural encoding of rudimentary, but not complex, sounds after peripheral deafferentation.

2. Cross-modal gain control in sensory thalamus

Michael Lohse, Johannes C. Dahmen, Victoria M. Bajo, Andrew J. King

Considerable attention has been focused on the involvement of cortical areas in multisensory processing. Although visual and somatosensory influences on auditory responses can arise via direct 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 nature of the multisensory interactions or the circuits involved in sensory thalamus. Using electrophysiology, 2-photon imaging of thalamocortical axons and optogenetic perturbations, we investigated multisensory interactions within different divisions of the medial geniculate body (MGB) and the circuits underlying these effects in mice.

Auditory responses were found to be reliably suppressed (~ 20 % reduction in firing rate across neurons) by somatosensory whisker inputs 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, with a subgroup of these medial neurons also being driven by whisker deflection.

Somatosensory suppressive influences were also found in auditory cortex, but absent in the auditory midbrain. In vivo 2-photon calcium imaging of auditory thalamic axons revealed that only the somatosensory inhibitory control of auditory thalamus is relayed to auditory cortex. Optogenetic activation of primary somatosensory cortex (S1) suggests that the pathway from S1 to thalamus is sufficient for mediating this crossmodal gain control. Using an intersectional optogenetic tagging approach based on projection target, cell type, and anatomical location, we are currently aiming to understand the intra-thalamic circuit contributions to the cross-modal gain control in auditory thalamus.

Our results suggest that sensory thalamus and surrounding circuits provide an important substrate for cross-modal gain control of auditory inputs.

3. Neural Selectivity for Music, Speech, and Song in Human Auditory Cortex

Sam V Norman-Haignere, Jenelle Feather, Peter Brunner, Anthony L Ritaccio, Josh H McDermott, Gerwin Schalk, Nancy Kanwisher

Music is ubiquitous across human cultures. How is it represented in the brain? fMRI studies have reported that cortical responses to music overlap with responses to speech and language, suggesting that music co-opts mechanisms adapted for other functions. However, we have previously found that when fMRI voxel responses are modeled as the weighted sum of responses from multiple neural populations, distinct selectivities for music and speech emerge. This finding suggests that the apparent overlap reported in prior studies is due to the coarse nature of fMRI, which blurs responses from nearby neural populations. To test this hypothesis, we measured cortical responses to a diverse set of natural sounds with human electrocorticography (ECoG), which has high spatial and temporal resolution, and which can sample from relatively large regions of the cortex. We observed clear selectivity for speech in some electrodes, and for music in others, validating our prior fMRI findings. Unexpectedly, we also observed electrodes that responded primarily to music with vocals (i.e. singing), and whose response could not be explained as purely a sum of music and speech selectivity. All category-selective responses developed quickly (within 200 to 500 ms of stimulus onset), and could not be explained by standard acoustic features. Music and song-selective responses were most prominent in anterior regions of the superior temporal gyrus (although a more posterior music-selective response was also observed), while speech selectivity was most prominent in the middle STG. These findings reveal that music and speech have distinct representations in the brain, but also that music itself is processed via multiple neural populations, one specific to the analysis of singing.

4. Processing of reverberant sound in the auditory system

Aleksandar Z. Ivanov, Benjamin D. B. Willmore, Kerry M. M. Walker, Andrew J. King, Nicol S. Harper

Every natural sound is accompanied by many delayed and distorted copies of itself (echoes or reverb). Unless the environment is very echoic, our brains cope effectively with reverberation. In contrast, reverberation can cause severe difficulties for speech recognition algorithms and hearing-impaired people. How might the healthy auditory system cope so well with reverberation? A major feature of auditory neurons is their ability to adapt to the sound statistics such as the mean or variance of the sound level (Dean et al. 2005, Rabinowitz et al. 2011). We posit that such adaptive phenomena could reduce the difficulties associated with reverberation. To test this hypothesis, we used a large data set of anechoic natural sounds and a room simulator to generate reverberant sounds. Both anechoic and reverberant sounds were passed through a model of the cochlea to produce 'cochleagrams'. We then trained a linear model to find the best mapping from the reverberant to the anechoic cochleagrams. The transformation learned by the units in this normative model displayed similar characteristics to the spectro-temporal receptive fields (STRFs) of auditory neurons, such as narrow frequency tuning and lagging inhibition similar to that observed in auditory neurons (Dean et al. 2008). For each model unit we measured a time constant of adaptation from this inhibition. The units showed time constants that were frequency-dependent, consistent with previous experimental data (Dean et al. 2008). The model also provided new predictions which we tested experimentally. First, the model's time constants increased with the amount of reverberation. Second, the ratio of inhibition to excitation in model STRFs increased with the amount of reverb. We tested these predictions in the auditory cortex of anaesthetised ferrets by recording responses extracellularly using Neuropixels probes. Our data suggest that the inhibitory time constants of real neurons, and the ratio of inhibition to excitation, both increase with the amount of reverb, similar to the model predictions. These findings suggest that the adaptive (Rabinowitz et al. 2011) and meta-adaptive (Robinson et al., 2008) properties of auditory neurons may play an important role in the processing of reverberant sounds.

5. Distinct codes as a substrate for causal inference in primate superior colliculus neurons

Jeff Mohl, Surya Tokdar, Jennifer Groh

Perceptual experience depends on combining information across multiple senses. For instance, auditory and visual information can be combined to produce better estimates of the spatial location of an object or event. However, cues should only be combined when they originate from the same source, and otherwise kept separate. This requires that the brain perform a causal inference in order to correctly combine (or not) multisensory information. It is unknown how and where this inference occurs.

We created a multisensory localization paradigm for primates which requires them to report the location of visual, auditory, and paired cues in both single (integrated) and multiple (segregated) target locations. This behavior was performed while we recorded single unit activity from a multisensory brain region, the superior colliculus (SC). Behavioral modeling shows that monkeys, like humans, are able to shift between integrating and segregating audio-visual information depending on target separation in a manner consistent with an approximation of optimal causal inference.

Contrary to previous studies that focus primarily on the “integrated” sensory condition, recorded SC neurons display a diversity of responses to combined stimuli. Most responses are best described by non-linear computations (such as a mixture of winner-take-all model) consistent with separated representations. Some neurons reflect approximate summation of unisensory responses (canonical multisensory enhancement), but these are much less common in our sample. These results suggest functionally distinct subpopulations, with some neurons representing stimuli as discrete (and separable) while others combine them into a single representation. Interactions between these subpopulations could provide a neural substrate for causal inference seen at the behavioral level.

6. Behavior state-dependence of correlated neural population activity in ferret primary auditory cortex

Charles Heller, Daniela Saderi, Zachary Schwartz, Stephen David

Sound-evoked spiking responses in primary auditory cortex (A1) neurons are often variable across repeated presentations of identical acoustic stimuli. Simultaneous recordings of multiple neurons in A1 have shown that some trial to trial fluctuations are correlated across the population. Moreover, the average magnitude of these spike count correlations (r_{sc}) changes with behavioral state, suggesting that they reflect underlying neural computations supporting acoustic perception and behavior. However, the circuitry underlying changes in r_{sc} and its dependence on behavioral state are not well understood. To address these questions, we recorded neural activity from A1 of awake, behaving ferrets using a linear 64-channel multi-electrode array and used pupillometry to simultaneously measure arousal. The linear array provided access to several (5-20) closely spaced single units within a single cortical column. We found that, on average, spike count correlations decreased during periods of active task engagement and during periods of high arousal. To investigate the mechanistic origin of arousal dependent variability, we modeled each neuron's response by scaling its stimulus response (PSTH) by pupil size. Thus, the model treated arousal as a linear, common input to primary auditory cortex. This analysis revealed that much of the arousal-dependent changes in spike count correlations are due to shared gain changes across pairs of neurons. In other words, when arousal-dependent gain changes were regressed out of single neuron responses, spike count correlations decreased across the population and no apparent difference in magnitude between high and low arousal state remained. Strikingly, this was not true for putative fast-spiking (FS) inhibitory neuron pairs. After accounting for the often observed arousal-dependent gain changes in FS neurons, the magnitude of spike count correlations in these pairs remained dependent on pupil size. Collectively these analyses suggest that pupil size reflects multiple, cell-type specific processes in primary auditory cortex.

7. Evaluating the generality of deep neural networks as a model of human hearing: Comparison with a large set of psychophysical and neural experiments

Alexander J.E. Kell, Erica N. Shook, Josh H. McDermott

Deep neural networks yield state-of-the-art performance on a variety of real-world perceptual tasks, and recent work has illustrated their utility as predictive models in sensory neuroscience. For instance, we recently found that a deep neural network optimized for speech and music recognition replicated human auditory behavioral characteristics on those tasks and predicted fMRI responses to a broad set of natural sounds better than other available models (Kell et al., 2018). However, deep networks are nonetheless incomplete models of sensory systems in many respects. For example, they suffer from “adversarial” examples to which human perception appears robust and likely require far more labeled training data than biological organisms receive over development. The extent to which deep learning can be used to further improve sensory models thus remains unclear. To help address this question, we examined whether such networks recapitulate the results of a wide range of psychophysical and neural experiments in humans.

We optimized a deep network for speech recognition in noise, training the network to map cochleagrams of speech excerpts to one of 587 word labels. To make the task more difficult and realistic, the speech clips were embedded in a variety of background noise sources (e.g., recordings of a bustling coffee shop, speech babble, etc.). We then measured the network’s generalization to a variety of classic stimulus alterations, including: (1) noise-vocoded speech (Shannon, 1995), (2) time-compressed and -dilated speech, (3) sine-wave speech (Remez et al., 1981), (4) locally reversed speech (Saberri & Perrott, 1999), and (5) speech in real-world reverberation (Traer & McDermott, 2016). We also tested whether the network used harmonicity as a cue to segregate speech from background noise, using resynthesized inharmonic speech (Popham et al., 2018).

In addition to these psychophysical effects, we simulated a number of recent fMRI experiments on the network. These included: deriving response “components” from the network units’ responses to natural sounds (Norman-Haignere et al., 2015), as well as measuring responses of each network layer to “quilted” speech (Overath et al., 2015) and “texturized” speech (McDermott and Simoncelli, 2011). Finally, we estimated spectrotemporal receptive fields (STRFs) for hidden units in different network layers.

The network recapitulated many psychophysical and neural effects previously observed in humans, but also exhibited notable discrepancies. These results extend existing evaluations of deep networks as models of audition, by providing a number of compelling matches to human behavior and neural responses. However, we also observed some failure modes on which future models can improve.

8. Enhanced ability of detecting rat vocalization-in-noise by sound exposure during a critical period

Natsumi Homma, Christoph Schreiner

During critical periods, receptive fields develop to adapt to the sensory environment by forming neural circuits that optimally process relevant stimuli. In humans, this developmental stage is essential for acquiring language and better sound processing. It has been shown that sound exposure during a critical period dynamically altered the receptive field properties in the auditory cortex such as frequency tuning, tuning bandwidth, or temporal resolution (de Villers-Sadani and Merzenich, 2011). However, the relationship between altered neural coding and perceptual abilities is yet largely unknown. In this study, we tested the hypothesis that exposure to moderate levels of structured background noise during the critical period enhances the ability of adult animals to process vocalization-in-noise.

Sprague-Dawley rat pups were raised in moderate noises (~60 dB SPL) of different spectro-temporal statistics during their auditory critical period (P6-45). Once these animals reached adulthood, they were trained to detect vocalizations presented in these noises using a Go/No-go behavioral paradigm and compared to unexposed animals. The sensitivity index (d') was calculated to evaluate the effect of different noise statistics on their ability of detecting vocalizations. Similarly, we investigated primary auditory cortical neuron responses to the same stimulus conditions in normal, exposed and trained animals. Neuronal discriminability was quantified using Euclidian distance-based spike train classifiers.

The noise exposure enhanced, in adult animals, their behavioral performance of detecting rat vocalizations in background noise. Improvement appeared to depend on stimulus statistics used for noise exposure. In addition, cortical signal encoding of vocalizations improved noise-exposed animals accompanied by specific shifts in receptive field properties compared to unexposed animals.

The results support the idea that maturational noise exposure can improve cortical receptive field properties best suited for information extraction in noisy environment thus reducing the impact of background noise masking and helping the animals to perceptually segregate signals from noise background.

9. Auditory cortex dependent reprogramming of an innate maternal behaviour

Alex Dunlap, Robert Liu

How sound information is transformed into appropriate behavioral responses is a central question in auditory neuroscience. Laboratory paradigms that operantly train animals to respond to target sounds have been useful in elucidating this question for sounds with nonsocial meanings. In social contexts, a promising paradigm for investigating how this occurs is the mouse maternal communication model, where through interacting with pups, mothers and cocaring mice come to preferentially approach ultrasonic isolation vocalizations to retrieve pups. While the retrieval act itself is instinctive in mice, and can be performed without prior pup experience, the degree of sensory flexibility in how this behavior can be triggered by novel sounds is not well understood. Presumably, the ability to preferentially approach cues that are predictive of infants would be highly adaptive for mothers, but how such new pup-associated stimuli would become integrated into maternal motor programs is unclear. Here, we use a pup reinforcer-based training paradigm to investigate whether female mice can incorporate a new auditory cue to guide their maternal motor program for pup retrieval. Trials begin with mice at its nest in the base of a T-maze. Two speakers are placed at either ends of the T's two arms, and a synthetic, amplitude-modulated noise stimulus is played from one of the two speakers to indicate which arm the mouse should enter to be given a pup. Pups, which are initially held outside the T-maze, are placed at the end of the arm associated with playback only after the mouse enters this "Correct" arm. The mouse then retrieves the pup back to the nest. If the mouse initially chooses the incorrect arm, they can subsequently investigate the other arm and receive the pup then without any punishments. We found that in this paradigm, mice have an innate strategy on the first day where on ~90% of the trials they return to the last arm where they previously received a pup on the last trial. In contrast, they correctly enter the playback arm on only ~45% of the trials on the first day, suggesting they do not use the novel sound to guide retrieval. Over the course of 8 days of training, the initial strategy shifts, wherein performance based on choosing the playback arm increases to ~80% of trials, while a strategy based on the last rewarded location drops to ~50%. After reversible, bilateral inactivation of auditory cortex with muscimol, we find that performance as measured by their initial strategy significantly increases ($n=6$, $p<0.05$), and performance based on sound playback drops significantly ($n=6$, $p<0.05$). These results indicate that auditory cortex is required for using a learned, pup-associated sound to overcome an innate strategy for guiding pup approach. Our paradigm allows investigating how downstream brain areas may be "reprogrammed" to utilize auditory cortical representations of pup-associated cues.

10. Increased cortical gain facilitates the detection of targets in noise.

Chris Angeloni, Maria Geffen

To function in highly variable acoustic contexts, it is necessary to adapt to persistent stimulus features to reliably differentiate sounds within the statistical regime of the current environment. A recent body of work demonstrated that neurons in auditory cortex (AC) adjust their gain to match the dynamic range of their spiking responses to the spectrotemporal characteristics of the stimulus (Rabinowitz et al., *Neuron*, 2011; Natan et al., *Cereb. Cortex*, 2017). Adaptive control of neural gain benefits loudness discrimination in environments with varying dynamic range (ie. different contrasts), but the effect of gain adaptation on target-in-noise detection is unknown. Here, we tested whether and how gain adaptation to different noise environments shapes neural and behavioral responses to embedded targets and examined the temporal dynamics of this process. To address these questions, we manipulated the contrast of a noisy background and embedded broadband targets at different temporal offsets relative to a change in contrast. Then, we recorded neural activity in AC of mice performing a target-in-noise detection task. In the neural population and behavior, target detection changed as predicted by gain adaptation to each noise environment, such that detection was facilitated by high gain and hindered by low gain. These observed behavioral and neural results were predicted by neurons simulated in a linear-nonlinear model, in which the gain of the neural nonlinearity was parametrically adapted over time. Our results suggest that gain adaptation, while useful for conserving discriminability across a range of stimulus contexts, can either facilitate or diminish target detection, depending on the statistics of the noise background.

11. Modulation of parietal alpha power reflects contributions of pitch cues to auditory spatial selective attention

Lia Bonacci, Barbara Shinn-Cunningham

In order to navigate complex scenes, an individual must select target stimuli while suppressing distractors. If the location of the target is known, this can be performed using top-down spatial attention. Neural correlates of spatial attention have been found using electroencephalography (EEG), namely through measurement of event-related potentials (ERP) and alpha (8-14 Hz) power. Increased ERP amplitudes from auditory cortex reflect enhancement of stimuli in the attended location. Increased parietal alpha power ipsilateral to the attended location is associated with suppression of distractors in the ignored location. While alpha oscillations have been studied extensively in vision, their role in auditory spatial attention is less clear. Furthermore, top-down spatial attention may not persist over the course of an auditory stream if pitch also differentiates target from distractors. We recorded EEG while subjects performed an auditory spatial attention task. Three melodies were presented simultaneously from different directions—left, right, and center—using interaural time differences of -100 mus, +100 mus, and 0 mus, respectively. Notes in each melody changed pitch over time, with contours that were rising, falling, or zigzagging. Subjects were cued to attend either the left or right melody and report its pitch contour. The center melody was always ignored. Experimental blocks alternated between two conditions that differed in the pitch separation of the competing melodies: one where the separation was large (~10 semitones) and one where it was small (~1 semitone). Passive trials, in which subjects ignored all stimuli and withheld a response, were also included. N1 amplitudes and alpha power were measured from the recorded EEG. In frontocentral sensors, N1s for a given note were larger when the note was attended than when it was ignored. Alpha power over parietal sensors varied with spatial attention focus. Compared to the passive condition, alpha power was larger in both left and right parietal sensors when subjects were cued to attend the left melody; when subjects were cued to attend the right, alpha power was larger over right parietal sensors. Importantly, while spatial attention focus changed alpha lateralization similarly in the two pitch conditions, this modulation was stronger when pitch differences were small. These results suggest that alpha modulation reflects suppression during auditory spatial selective attention. When pitch cues are strong, parietal alpha modulation is weak, likely reflecting the fact that pitch differences can be used to help focus attention.

12. Competing Sounds reveal Spatial Configuration Sensitive Neurons in Mouse Primary Auditory Cortex

Howard Gritton, Kevin Lou, Jun Ma, Xue Han and Kamal Sen

Everyday auditory scenes typically contain multiple sound sources at different spatial locations. Although many previous studies have explored cortical spatial processing of single sound sources, cortical processing of competing sound sources remains poorly understood. Here, we report auditory cortical neurons in mouse A1 with spatial response properties that are well suited for segregating competing sounds from different spatial locations. We found that auditory cortical neurons are broadly tuned to single “target” sound sources from different spatial locations; however, when the target is presented at the same time as a competing “masker” from a different location, cortical neurons sharpen their spatial tuning. Specifically, we quantified neural discrimination performance over a spatial grid of different combinations of target and masker locations and found that cortical neurons display “hotspots” of high performance at particular positions on the spatial grid. Thus, such neurons are sensitive to the spatial configuration of the target and the masker. We find that individual sites show diverse spatial grids, and the cortical population allows for a good representation of all configurations, for spatially separated target and maskers. Furthermore, we find that fine temporal structure in neural responses is important for achieving best neural discrimination performance for most sites.

Similar findings have been previously reported for cortical level neurons in songbirds. Songbirds and mice have different frequency ranges of hearing, and different peripheral representations for frequency dependent acoustic cues, e.g., interaural time-difference (ITD), and interaural level difference (ILD). Our findings therefore suggest that the emergence of general cortical spatial representations, despite different peripheral representations of acoustic cues, is a biologically conserved property of neural sound processing

13. Predictive auditory sequence learning modulates inter-regional oscillatory coupling in human intracranial recordings

Yukiko Kikuchi, Christopher K. Kovach, Ryan Calmus, Phillip E. Gander, Ariane E. Rhone, Kirill V. Nourski, Hiroto Kawasaki, Timothy D. Griffiths, Matthew A. Howard III & Christopher I. Petkov

Dynamic environments generate sequences of sensory events that affect brain oscillations at different frequencies. Moreover, feed-forward and feed-back interactions are evident in oscillatory phase and frequency interactions, yet little is known about how learning-related plasticity changes oscillatory dynamics in neural networks. Using intracranial recordings from both human and monkey auditory cortex, we had previously reported phase-amplitude coupling between low-frequency phase and gamma amplitude in auditory cortex after the learning of temporal relationships in sequences of speech sounds drawn from an Artificial Grammar (AG; Kikuchi et al., PLoS Biology, 2017, 15(4):e200219). In the monkeys as an animal model, the prior results pointed to oscillatory influences on auditory cortex neurons emanating from other brain regions. Here, we sought to identify the sources of these influences using the extensive coverage with depth and grid recording electrodes offered by human intracranial recordings in epilepsy patients being monitored for surgery. We used the same AG learning paradigm as in the prior study, where participants were first exposed to representative rule-based sequences of nonsense words. In the subsequent testing phase, we presented the participants with sequences that were either consistent with the AG or contained specific violations of learned ordering relationships. Inter-regional phase interactions were measured by phase angle differences between recording sites weighted by the evolution of the transitional probabilities over all previously encountered sequences. Sequencing-related modulation of oscillatory phase coupling was observed between auditory cortex (Heschl's gyrus, HG) and a number of sites, including the hippocampus, lateral superior temporal gyrus, anterior temporal lobe and inferior frontal gyrus (IFG). Further analyses showed that frequency-specific phase coupling tracked the changes in transitional probabilities within three oscillatory bands (theta: 3-8 Hz, beta: 15-29 Hz, gamma: 40-100 Hz). Moreover, during initial learning, HG interacted with the hippocampus and IFG. A more extensive network was involved after the learning phase, with cortical network modulation in response to violations of the learned sequencing relationships involving a broader set of regions. Oscillatory phase-based coupling in human intracranial recordings provides insight into inter-regional neural interactions at different stages of predictive sequence learning, which refines hypotheses for further testing using monkeys as a model system.

14. The eardrums report both eye position and eye movement: the full EMREO map

David LK Murphy, Rachel Landrum, Cole D. Jenson, Stephanie Schlebusch, David W. Smith, Christopher A. Shera, Cynthia D. King, Jennifer M. Groh

After every eye movement, the brain must realign the visual and auditory reference frames in order to co-locate sights and sounds. Exactly where, when, and how such visual-auditory spatial integrations occur is not fully understood. We recently discovered that the auditory periphery receives a signal from the brain that causes the eardrum to oscillate beginning a few milliseconds before saccades and continuing until well into ensuing periods of fixation (Gruters, Murphy et al PNAS 2018). Information about at least the horizontal direction and length of the saccade are reflected in the phase and magnitude of these eye movement-related eardrum oscillations (EMREO). Here, we sought to assess the full spatial characteristics of this signal in two dimensions, including the relative contributions of eye displacement vs. absolute (initial or final) eye position.

We found that EMREOs depend on both horizontal and vertical dimensions, and are impacted by both eye displacement and absolute (initial or final) position. In toto, the vertical and horizontal saccade parameters (displacement and initial position) produce distinct oscillatory components, suggesting that EMREO uniquely represent combinations of saccade parameters. These results demonstrate that detailed information about the relationship between visual and auditory reference frames is present in the earliest stage of the auditory pathway. Future work delving into the relationship between EMREO and the transduction of incoming sounds will be needed to ascertain their effects on the processing of auditory spatial locations in relation to the visual scene.

15. Reproducibility of eardrum movements accompanying saccades: implications for clinical testing

Stephanie Schlebusch, David L.K. Murphy, Cynthia King, David M. Kaylie, Jennifer M. Groh

The visual and auditory systems play mutually reinforcing roles in perception. This is apparent in clinical research and treatment of deficits affecting one of these modalities. Although where in the brain this connection occurs is unknown, we have recently reported an oscillation of the eardrum that accompanies saccades. The phase and magnitude of these eye movement-related eardrum oscillations or EMREOs covary with the direction and magnitude of the saccade (Gruters, Murphy et al. PNAS 2018). This finding suggests that the connection between the visual and auditory systems may begin as early as the auditory periphery.

We sought to determine how reproducible this oscillatory signal is in an individual primate subject across days and how much data is required to reliably estimate the EMREO of a new subject. Information about reproducibility can help lay the groundwork for extending these findings to the clinical setting.

We tested one rhesus monkey repeatedly in ~30 one-hour sessions over a two month period. The monkey was head-restrained and made saccades spontaneously in a dark room. Eardrum oscillations were recorded using microphones placed in the ear canals of both ears and eye movements were tracked with a video eye tracker (1000 Hz sampling rate).

We report that the EMREO in a primate can be assessed satisfactorily after about 100 well-identified saccades spanning a range of directions and amplitudes. This currently takes about five minutes of data collection, and would likely be faster in human subjects performing a visual saccade task. The assessments obtained with spontaneous saccades are stable and repeatable across days over a two month period. Together, these findings bode well for the feasibility of testing the EMREO signal in the clinical setting.

16. Auditory Attention Decoding: What Anatomical Locations and Neural Frequency Bands Contribute?

James O'Sullivan, Jose Herrero, Elliot Smith, Sameer A. Sheth, Guy M. McKhann, Ashesh D. Mehta, Nima Mesgarani

Decoding an attended speaker from neural recordings (termed auditory attention decoding; AAD) has many applications. The most pertinent is probably the development of a cognitively controlled hearing aid that can automatically track and amplify an attended speaker. Such devices will likely be limited to either non-invasive or minimally invasive neural recordings. Non-invasive recordings such as electroencephalography (EEG) can typically only record from low frequency (LF; < 50Hz) neural data, and have relatively poor spatial resolution. However, multiple electrodes can be used to target cortical areas using source-localization signal processing strategies. Minimally invasive approaches can place a restricted number of electrodes over specific cortical areas, and can also record from higher neural frequencies (<200Hz). In both cases, knowledge of the anatomical locations and neural frequency bands that contribute to AAD is crucial.

To investigate, we used an invasive recording methodology known as electrocorticography (ECoG) that can record from both low and high frequency (HF; < 200Hz) neural data, and can also localize neural activity to within ~3mm from both deep and surface brain regions, spanning the full extent of auditory cortex. We show that both LF and HF data, as well as deep and surface regions, can be used to decode attention. However, we found a dichotomy between the combinations of frequency band and anatomical location that could be used: when using HF data, the anatomical region that produced the most robust encoding of attended speech was superior temporal gyrus (STG; a surface brain region). Conversely, LF data was the best at decoding attention in Heschl's gyrus (HG; a deep brain region). Both of these combinations (LF data in HG, and HF data in STG) provided similar results in terms of decoding speed and accuracy. These results provide the first extensive exploration of the neural frequency bands and anatomical locations that contribute to AAD, and will inform future work on the development of cognitively controlled hearing aids.

17. Integration of cross-modal information enhances auditory gap detection performance

Anna-Katharina R. Bauer, Martin G. Bleichner, Sylvain Baillet, Stefan Debener

Our sensory systems often receive asynchronous but related input, so that visible events often precede and cause subsequent sounds. In two independent experiments, using electroencephalography (EEG; N = 28) or magnetoencephalography (MEG; N = 12) respectively, we investigated whether visual rhythmic stimulation modulates subsequent auditory cortex oscillatory activity by reorganising the phase of auditory cortex oscillation. Listeners were presented with auditory-only and visual-auditory stimuli and participants had to detect short silent gaps, that were uniformly distributed with respect to the phase of a 3 Hz frequency-modulated tone. In the visual-auditory condition, the visual stimulation consisted of a Gaussian pulsating circle, which preceded the auditory tone and was presented either in or out of phase. Gap detection performance increased for the visual-auditory conditions relative to the auditory-only condition and accuracies were modulated by the stimulus phase. Analysis of the power spectral density revealed spectral peaks at 3 Hz and the 6 Hz harmonic at the sensor (EEG) and the source level (MEG). Further, analysis of inter-trial phase coherence revealed a single peak in the 3 Hz frequency band for both visual and auditory stimulation (EEG & MEG). Cross-modal spectral analysis revealed enhanced power in the auditory cortex during visual stimulation (MEG), suggesting that the auditory cortex prepares for the upcoming visual stimulation. The cross-modal temporal dynamics are characterised by a sharp increase of the stimulus-brain coupling for both visual and auditory regions of interest (MEG). In two studies, we showed that cross-modal phase entrainment leads to enhanced gap detection performance and neural entrainment effects, suggesting that visual rhythmic stimulation has beneficial effects on auditory perception.

18. Multiscale Calcium Imaging of Auditory Cortex in Awake Marmoset Monkeys

Xindong Song, Yueqi Guo, Xiaoqin Wang

The common marmoset (*Callithrix jacchus*), a highly vocal New World monkey species, has emerged in recent years as a promising non-human primate model for neuroscience research. Because the marmoset brain is lissencephalic (smooth), nearly all cortical areas are accessible directly under the skull with optical imaging methods. Two-photon calcium imaging with genetically encoded calcium indicators (GECIs) has been previously implemented in marmoset somatosensory cortex. Two-photon imaging in the auditory cortex is technically challenging since a mechanically vibrating laser scanner can generate sounds that are audible to the animal and thus interferes with the experimental design. In the current study, we developed a flexible, agile, yet silent two-photon microscope based on an acousto-optical deflector (AOD) scanner. An optical window with a quarter-inch diameter was implanted over the auditory cortex of awake marmosets. A clear tonotopic structure can be observed through intrinsic signal imaging at both green and blue wavelengths. Multiple virus injections carrying GCaMP were made through the silicone-based optical window. A dual-virus strategy was used to separate controls over expression specificity and expression level. A clear macroscopic wide-field fluorescence response was observed starting 10 days after virus injections, at sound levels as low as 10 dB lower than hearing thresholds. The tonal response recorded from wide-field fluorescence imaging was consistent with intrinsic signal imaging results and was also limited mainly within the primary auditory cortex. By using complex sound stimuli, such as music or marmoset vocalization recordings, strong and widespread cortical responses ($\Delta F/F > 10\%$) can be evoked in both primary and secondary cortical areas. The silicone-based window can be replaced by a glass coverslip-based window. And a customized silent two-photon microscope was used to measure responses of individual neurons at a microscopic scale. The general response patterns within each two-photon field-of-view were consistent with wide-field imaging results. However, individual neurons' responses can be heterogeneous even for close-by neurons. The multiscale calcium imaging approach reported here thus provides a new experimental paradigm for functional mapping of the marmoset auditory cortex in the awake condition in a high throughput way over conventionally electrophysiology methods.

19. Decoding sound texture identity via statistics of neuron ensembles

Xiu Zhai, Mina Sadeghi, Fatemeh Khatami, Heather L Read, Ian Stevenson, Monty A Escabi

Although pure tones and noise stimuli have been used extensively in auditory neuroscience, there is growing evidence that the higher-order sound statistics, such as the correlations between frequency channels or the sound modulation spectrum, play a key role in sound recognition and perception. How the brain encodes such sound statistics and utilizes them for sound identification is largely unknown. Here we examine the responses of neuron ensembles in the auditory midbrain (inferior colliculus) of unanesthetized rabbits listening to natural sound textures using 16 and 32 channel neural recording arrays. Sound textures, such as running water, fire, wind, and speech babble have complex, but homogeneous, higher-order statistics, and are perceptually salient. We use texture synthesis (McDermott & Simoncelli 2011) to manipulate the different statistics of natural sound textures and explore their neural representation. Five texture sounds (running water, bird chorus, crackling fire, rattling snake and crowd noise) were manipulated by progressively incorporating statistical structure from the power spectrum, amplitude marginals, modulation spectrum, and the sound correlation structure. We demonstrate that correlated firing between frequency organized recording sites are modulated by each of the sound tested statistics and that these neural correlations can be used to decode and identify sound textures. Specifically, stimulus-driven spectro-temporal correlations were measured across the frequency organized recording array and a minimum distance classifier was applied to the ensemble correlation activity to identify the delivered sounds. For the original texture sounds, the classifier was able to decode and identify the original sound approaching near perfect accuracy (~90%). Spectral correlations between recording locations had slightly higher performance and were somewhat more informative than temporal correlations. The performance of the classifier improved as additional statistics were added to the synthetic variants and approached the performance for the original sounds when the full set of statistics was included (~80% accuracy). Finally, the decoding accuracy improved with sound duration with evidence accumulation times in the order of approximately 1 sec, mirroring human trends. These findings suggest that coordinated firing in auditory midbrain ensembles provide a statistical signature that may contribute to perception and recognition of texture sounds (supported by NIDCD R01DC015138).

20. Multiplexed encoding of sensory and reinforcement cues in the cholinergic basal forebrain supports associative learning and cortical plasticity

Blaise Robert, Wei Guo, Daniel B. Polley

The rodent basal forebrain is studded with dense clusters of cholinergic neurons. Artificial stimulation of these areas can produce striking plasticity in sensory cortex, but an organizing schema for understanding the natural activators of cholinergic basal forebrain units has proven elusive due to the technical difficulty of recording from genetically defined cell types across this distributed deep brain network in behaving animals. Here, we describe an optogenetic antidromic phototagging approach to isolate single cholinergic units in Nucleus Basalis (NB) that project to the auditory cortex (ACx) and characterize their response properties and learning-related plasticity in awake, head-fixed mice.

We found that phototagged cholinergic NB --> ACx units (ChACx) and neighboring non-cholinergic NB units (NChACx) had robust, short-latency responses to meaningless, unconditioned auditory stimuli such as tones or noise bursts. ChACx and NChACx units also showed robust responses to unpleasant stimuli, such as air puffs directed at the face. Because NB unit spike trains multiplex neutral sensory stimuli and unconditioned behavioral reinforcers, we reasoned that they could play a critical role in linking conditioned stimuli with temporally delayed (i.e., distal) reinforcement cues. We tested this by pairing pure tone stimuli (CS+) followed by air puffs (US) 5s later and confirmed behavioral evidence of associative trace learning. We observed a rapid tone-specific plasticity in the frequency receptive fields of ACx units and ChACx units, but not NChACx units. Paired recordings from ACx and NB over the course of conditioning revealed a strong increase in gamma band coherence that parallels a selective potentiation of spike-evoked L2/3 local network activity from ChACx units during the initial CS and US pairings.

These findings identify a new role for cholinergic units in the caudolateral extreme of the basal forebrain (Nucleus Basalis) in driving cortical plasticity that supports learned associations between sensory conditioned stimuli and distal, temporally delayed reinforcement cues. Our ongoing studies utilize deep-brain imaging methods to identify regional differences among cholinergic basal forebrain neurons to sensory stimuli, behavioral reinforcement valence, and the encoding of sensory and reinforcement error signals. By discerning underlying rules for activating these distributed deep brain neuromodulatory neurons, we hope to develop new approaches to efficiently transform neocortical sensory representations.

21. Functional organization of human perisylvian cortex in response to speech

Bahar Khalighinejad, Sam Norman-Haignere, Jose Herrero, Ashesh Mehta, Nima Mesgarani

Auditory cortex is thought to be organized into different cortical fields, each of which encodes multiple acoustic properties. Functional characterization of human cortical fields has largely depended on non-invasive brain imaging techniques such as fMRI. However, because fMRI indirectly measures neural responses via hemodynamic activity, it has limited temporal resolution, making it difficult to study the neural encoding of speech features which vary on the timescale of milliseconds. Thus, here we used direct intracranial recordings from 12 human subjects that were implanted with depth and grid electrodes.

We recorded neural data in response to 40 minutes of natural speech and other commonly heard sounds. Using 350 electrodes in primary and secondary auditory cortices, we created seven cortical maps based on tuning for each of seven different acoustic features: frequency, latency, temporal modulation, spectral modulation, phonetic features, speaker features, and speech-specificity. The frequency and latency preferences were calculated from spectro-temporal receptive fields (STRF), estimated from each electrode. The temporal and spectral modulation maps were calculated by filtering the acoustic spectrograms through different cochlear filters and then by selecting the cochlear filters that best represent the temporal modulation preference and spectral modulation preference of the neural data.

Consistent with prior fMRI studies, we found that topographic maps of frequency preference, latency, temporal modulation, and spectral modulation were dependent on two axes of medial-lateral and posterior-anterior direction in human auditory cortex. But contrary to prior fMRI studies, we did not find any specialized region for analyzing phonetic features, meaning that all of the four regions of medial Heschl's gyrus (HG), lateral Heschl's gyrus, planum temporale, and superior temporal gyrus (STG) encode the distinctive acoustic features of phonemes, and there was not a significant difference between their preferred phonetic feature. In addition, we found that information about speaker identity was better encoded in early auditory areas such as HG in comparison with secondary areas such as STG. Finally, we found that electrodes in STG selectively responded more to human speech compared to nonspeech sounds, such as animal vocalization, tones and music, confirming prior reports of speech-selectivity based on fMRI. Together, these findings advance our knowledge of the representational and functional organization of human auditory cortex and pave the way toward more complete models of cortical speech processing in the human brain.

22. Mechanisms of information processing by coordinated neuronal ensembles in the primary auditory cortex

Jermyn Z See, Craig A Atencio, Natsumi Y Homma, Vikaas S Sohal, Christoph E Schreiner

How individual neurons work together to encode sensory information and influence behavior remains one of the fundamental questions in sensory neuroscience. 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, assuming 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 analyses.

We recently showed that coordinated neuronal ensembles (cNEs, i.e. groups of neurons with reliable synchronous activity) are stable functional constructs that may represent the basic unit of information processing in AI (See et al., 2018). To further assess the functional properties of cNEs in AI, we performed dense extracellular recordings in rat AI while presenting dynamic, broadband stimuli, identified cNEs using dimensionality reduction techniques, and isolated cNE events, using them to assess spectrotemporal information processing. We found that there were different types of cNEs – some contain and even improve stimulus information over that of member neurons, while others seem to be agnostic to the presented stimuli and may represent convergence of top-down inputs.

These findings challenge the classical idea that AI neurons produce a homogeneous set of spikes that may be equally weighted to estimate a neuron's receptive field. Instead, individual spikes of AI neurons may represent different pieces of information, from bottom-up auditory information to top-down information that can be independent of auditory processing aspects. Therefore, by taking into account the stimulus preferences or lack of stimulus preference associated with each cNE, we may gain a more complete understanding of multiple aspects of information processing in AI networks.

23. Investigate what makes it “new” in the old-plus-new strategy of auditory scene analysis in the auditory cortex of marmoset monkeys

Yi Zhou, Jonas Braasch

Auditory Scene Analysis (ASA) is an umbrella term that describes the auditory system’s remarkable ability to decompose signal-based acoustic features from a complex sound mixture and regroup them into perceptually relevant auditory objects and streams. A fundamental concept of ASA is the old-plus-new strategy, which states that the auditory system correctly interprets the sudden development of new elements in an acoustic mixture as additional, distinct elements added to an existing continuing sound source. Braasch and Hartung (2002) showed that human localization performance of a target sound (200-ms broadband noise) was worsened by a masker sound (different noise waveform) presented simultaneously from a fixed frontal location. However, the listeners had no difficulties localizing the target if the masker preceded the target by 200 milliseconds. Their model proposed that the correct target localization was possible because listeners could use information from the preceding part of the masker to perceptually separate target and masker using the old-plus-new strategy of ASA.

This study investigated the neural mechanisms of perceptual spatial release from masking and the way in which neuronal activities manifest the target spatial information in the presence of a competing masker sound. We used a similar spatial setup to the one used in the human study mentioned above. We collected single-unit activity from the auditory cortex of awake marmoset monkeys. We first measured the rate spatial tuning function of a neuron to a target sound (broadband noise or best-frequency tone) presented alone. We then compared how simultaneous or preceding noise or tonal masker altered the target-alone spatial tuning function. We observed that while the spiking activities of many cortical neurons show hemifield-field spatial tunings (front, back, left and right), the non-responsive regions are actively inhibited, suggesting cortical modulations of auditory spatial selectivity. We found that spatial tuning remains less affected with the preceding than simultaneous maskers, consistent with human studies. However, the extent of spatial release from masking is affected by the onset/sustained patterns of the masker responses and the extent of cortical inhibition/suppression in both target and masker responses. We propose a new model for the old-plus-new strategy in the spatial domain, which integrates two neural processes – adaption and suppression – to reveal the central mechanisms that make the target spatial information become “new” in a multi-source acoustic environment.

24. Mechanisms to Communicate in a marmoset 'Cocktail Party'

Vladimir Jovanovic, Cory T. Miller

Effective vocal communication relies on the capacity to parse meaningful information from an abundance of interfering noise in the environment. For humans, the challenge of attending to a single speaker in a noisy setting comprised of multiple other conspecifics engaged in conversation is commonly referred to as the 'Cocktail Party Problem'. Here we developed an innovative paradigm that simulates a naturally occurring marmoset cocktail party to examine the underlying perceptual mechanisms. We utilized our Virtual Monkey (VM) paradigm in which calls are broadcast at times that simulate the natural vocal behavior of the species. In this design, freely-moving subjects are presented with vocalizations from five VMs from different speakers in the test environment. Two pairs of these VMs (Distractors) engage in conversational exchanges only with each other. In contrast, the vocalizations emitted from the fifth VM (Target) occurred only in response to subject's vocalization. Given the acoustic complexity of this environment, we then asked whether subjects would be able to perceptually parse the Target VM's vocalizations amidst the frequent vocalizations from the Distractors. Analyses focused on three facets of subjects' vocal behavior relative to all VMs: Response Rate, Response Latency, and ratio of subject calls in a Conversation with a VM (i.e. multiple consecutive, reciprocal vocal exchanges). Results showed a consistent bias towards the Target VM. Subject response rate to Target is 30% higher, latency 3 sec faster, and calls in a conversation occurred 45% more often, compared to all other Distractors. These data suggest that, like humans, marmosets are able to effectively attend to a single speaker in an environment comprising high levels of acoustic interference from conspecifics. Current experiments aim to systematically test the perceptual mechanisms that support effective communication in a marmoset cocktail party.

25. Cross-Modal Representation of Individual Identity in Marmoset Hippocampal Neurons

Hristos Courellis, Cory Miller

The convergence of multi-modal sensory information in the hippocampus facilitates the internal representation of a complex external world. A human, for example, integrates information across multiple sensory modalities, such as their voice and face, to establish a cohesive representation of individual identity, a process supported by mechanisms in hippocampus. Here we sought to interrogate whether neurons in the hippocampus of an awake marmoset monkeys (*Callithrix jacchus*) exhibited mechanisms for individual identity. Like humans, this primate species routinely uses both visual and acoustic social signals to mediate their conspecific social interactions. We presented subjects with concurrent visual and auditory signals – faces and vocalizations - collected from their family members and unrelated animals in the colony. We hypothesized that single units in the marmoset hippocampus would be differentially responsive to these sensory signals when presented separately and presented concurrently. We further hypothesized that a sub-population of neurons would be sensitive to the congruence of these signals for representing the individual identity of the monkey. We implanted chronic electrodes into marmoset hippocampus using a structural-MRI guided surgical procedure. Marmosets were head-fixed before a digital display and a speaker to present both auditory and visual stimuli. Task structure involved a brief fixation after which either a vocalization, face, or combined audiovisual stimulus were presented. Audiovisual stimuli were further sub-divided into matched and mismatched stimulus identity. We found hippocampal single units to be significantly differentially responsive to audiovisual combinations of sensory stimulation. Furthermore, we identified a population of neurons sensitive to the congruence of identity between auditory and visual stimuli. These initial analyses suggest that neurons in marmoset hippocampus contribute to a cross-modal representation of individual identity.

26. Contribution of correlated cortical activity to identity-preserving changes in sounds.

Francisco A. Rodriguez Campos, Matthew Schaff, Brianna Karpowicz, Yale E. Cohen

Abstract: Our environment is filled with acoustic stimuli that our brains transform from low-level sensory responses into perceptual representations called auditory objects. These objects are the foundational building blocks of our auditory-perceptual world and are the computational result of the brain's capacity to detect, extract, segregate, and group the regularities in an acoustic environment, a process often referred to as "auditory scene analysis". Listeners are tolerant to substantial changes in the spectral and temporal features of auditory objects. For example, as a singer moves across a stage, we can recognize their unique voice and understand the song although the acoustic features vary dramatically with location. In order to enable such invariance, the hierarchy of cortical areas must develop neuronal representations of auditory objects that are increasingly independent of their naturally occurring transformations in time. However, the brain areas, neural mechanisms, and computations that provide the basis for a listener's tolerance to these identity-preserving changes are unknown.

To address the neuronal basis of invariant sound recognition, we conducted large-scale recordings (MicroProbes' 96-electrode Microwire Brush Array) from the auditory cortex (anterolateral belt [AL], middle-lateral belt [ML], and primary auditory cortex [A1]) 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 changes in the location of the stimuli and changes in 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 sound source as the original stimuli. We report how noise correlations vary with the identity-preserving transformations inflicted by our sound ensembles. In particular, we describe how the spiking activity and population activity in those three brain regions support representations of invariance

27. A modular high-density 294 channels μ ECoG system on macaque vIPFC for auditory cognitive decoding

Chia-Han Chiang, Jaejin Lee, Charles Wang, Ashley J. Williams, Yale Cohen, and Jonathan Viventi

Discerning sound is a complex neurocomputational problem because the stimuli of interest in the real world change simultaneously along multiple dimensions and are often mixed together with other environmental sounds. It is not well understood how the brain transforms a mixture of acoustic stimuli into distinct perceptual representations. There is broad consensus that auditory perceptual decisions are mediated by the ventral auditory pathway. At the later stages of this pathway, including the ventrolateral prefrontal cortex (vIPFC), neurons can encode decision outcomes. However, the spatiotemporal resolution of the neural code that underlies perception and cognition over this region is still unclear; and until recently, it was not possible to record from large numbers of brain sites simultaneously. We, therefore, developed a modular high-resolution electrode array system with long-term viability and used it to study the information that could be decoded from vIPFC micro-electrocorticographic (μ ECoG) signals during an auditory detection task.

We molded three separate μ ECoG arrays into one with silicone and implanted this high-density modular system in one adult male rhesus macaque. A custom 3D-printed titanium chamber was mounted on left hemisphere using MRI-based targeting (BrainSight software, Rogue Research). The molded 294-contact polyimide μ ECoG array (30 μ m total thickness; 229 μ m diameter gold contacts; 610 μ m contact spacing; 10.4 by 11 mm² sensing area) was implanted subdurally over vIPFC. Intan headstages and an OpenEphys recording system were used to record μ ECoG activity while the monkey participated in a “hearing-in-noise” task in which they reported hearing a “target” vocalization from a background “chorus” of vocalizations. We titrated task difficulty by varying sound level of the target vocalization, relative to the chorus.

We present a decoding analysis of the μ ECoG signal relationship to behavior with respect to spatial resolution, neural frequency bands, time, and cognitive/behavioral state. We found that individual channels varied significantly in their ability to decode various behavioral parameters (e.g., hits versus misses): some channels were at chance, whereas others performed significantly better than chance. When we considered those channels that performed better than chance, we found that, as we increased the number of channels, decoding performance improved. This improvement was seen across a variety of neural frequency bands.

28. Modulation of auditory cortical information processing by movement and VIP interneuron activation

James Bigelow, Jefferson DeKloe, Ryan Morrill, Andrea Hasenstaub

Information processing in sensory cortex is highly sensitive to contextual variables such as anesthetic state, arousal, and task engagement. Recent work in visual cortex (VCTX) has established that local circuitry responsible for extracting information from visual environment is highly sensitive to inputs originating from motor circuits activated during locomotion, finding that evoked firing rates and stimulus information increase when animals are engaged in movement. A specific inhibitory interneuron circuit appears to be critical for this change. Inhibitory interneurons expressing vasoactive intestinal peptide (VIP) are differentially activated during movement, which typically suppress other inhibitory interneurons, ultimately disinhibiting excitatory pyramidal cells. Although VIP activation has been observed during movement in somatosensory and auditory cortices, it remains unclear whether these activations similarly elevate evoked responses and stimulus information. The present study examined auditory cortical (ACTX) responses evoked by tone cloud stimuli in awake, headfixed mice during spontaneous movement and still conditions. To test the role of a VIP-mediated circuit in motor-related activity modulation, we crossed VIP-Cre mice with Ai32 mice to express channelrhodopsin in Cre-expressing VIP interneurons. VIP+ cells were optogenetically activated for half of the stimulus presentations, permitting independent analysis of the consequences of movement and VIP activation, as well as their intersection. Preliminary analyses suggest heterogeneous influences of both movement and VIP activation on ACTX responses. In contrast to VCTX, stimulus-evoked spike counts tend to decrease during movement in our neuron sample. Loss of spikes during movement appears to be associated with a reduction in total information (bits) as well as information efficiency (bits/spike). Consistent with the disinhibitory circuit observed in VCTX, VIP interneuron activation tends to elevate evoked firing rates in ACTX neurons. Importantly, however, the additional spikes produced by VIP activation seem to be unrelated to the stimuli, usually contributing zero additional information (bits), thus undermining information efficiency (bits/spike). The effects of simultaneous movement and VIP activation appear to sum linearly: although evoked spike counts during locomotion generally return to baseline levels with concurrent VIP activation, information remains abnormally low. Our findings raise intriguing possibilities about asymmetric consequences of motor circuit activation on information propagation in VCTX and ACTX.

29. Reorganization of cortical population neuronal activity following auditory fear conditioning

Katherine C. Wood, Richard Betzel, Danielle Bassett, Maria N. Geffen

Auditory perception relies on learning-driven neuronal plasticity within the auditory pathway. Here, we investigated how associative learning, differential auditory fear conditioning (DAFC), affects neuronal population responses to sounds in auditory cortex (AC). In DAFC, the subject is presented with two different frequency tones, one of which is paired with a foot-shock. Previously, we found that AC is required for expression of DAFC-driven changes in sound-frequency discrimination acuity (Aizenberg and Geffen, 2013) and that modulating inhibitory neuronal activity in AC leads to similar bi-directional changes in discrimination acuity (Aizenberg, 2015). However, how DAFC affects tone-evoked population neuronal activity remained unknown. We hypothesized that DAFC would drive changes in population tone-evoked neuronal activity corresponding to either an increase or a decrease in neurometric frequency discrimination acuity, as a function of fear learning specificity. To understand the transformation of sound representation in AC before and after DAFC we imaged calcium activity in hundreds of neurons simultaneously in AC of awake, head-fixed mice, tracking the same neurons over days under a two-photon microscope before and after two DAFC sessions. We quantified changes in tone frequency-dependent responses of individual neurons, as well as in population functional connectivity. DAFC drove heterogeneous changes in individual neuronal responses for either shock-paired or unpaired tone frequencies. At the same time, mean population neuronal response strength to tones across frequencies was preserved. However, neuronal responses to tones following DAFC became more consistent after DAFC. Neuronal populations formed clusters driven by correlated activity, neurons within clusters exhibit heterogeneous response patterns. The neuronal cluster structure changed between days in the absence of DAFC, but the network structure became more consistent over days following DAFC. These findings suggest that DAFC drives cortical population activity toward a more stable state.

30. Connectional modularity within the lateral cortex of the mouse inferior colliculus gives rise to partially segregated processing streams for auditory and multisensory information
Alexandria Lesicko, Daniel Llano

The lateral cortex of the inferior colliculus contains a network of modules characterized by dense staining 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, it is unknown whether the intrinsic connections in the lateral cortex also exhibit connectional modularity. In the present study, we sought to characterize the intrinsic inputs to excitatory (GAD-67-) and inhibitory (GAD-67+) neurons in both modular and extramodular regions of the lateral cortex. 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+ and GAD-67- cells in both regions were filled and 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. Morphological reconstructions of biocytin-filled cells revealed that the dendrites of neurons in the lateral cortex are largely confined to the domain (modular or extramodular) in which their cell body resides, with the exception of GABAergic modular cells. Photostimulation maps generated under synaptic blockade further support these results; direct stimulation of extramodular cells is elicited from extramodular sites, direct stimulation of non-GABAergic modular cells arises from modular sites, but direct stimulation of GABAergic modular cells can be driven from both domains. Pre-synaptic photostimulation and spatial analysis revealed that extramodular cells receive input almost exclusively from the extramodular domain, non-GABAergic modular cells receive mixed input from both domains, and GABAergic modular cells receive the majority of their input from the extramodular domain. Overall, these results indicate that there is a unidirectional flow of information within the lateral cortex, such that modular cells receive inputs from auditory-recipient (extramodular) and somatosensory-recipient (modular) areas of the lateral cortex, while extramodular cells only receive input from the extramodular domain. This modularity in the intrinsic connectivity may give rise to partially segregated processing streams for auditory and multisensory processing in the lateral cortex of the inferior colliculus.

31. How do neurons overcome developmental hearing loss-induced cellular deficits during auditory learning?

Todd Mowery, Nihaad Paraouty

Developmental hearing loss leads to deficits in the auditory perceptual abilities of children and these arise from peripheral and central changes along the auditory neuraxis. When peripheral deficits are treated through the resolution of ear canal blockage, surgery, hearing aids or cochlear implants, perceptual thresholds often return to control values. When tested on some non-perceptual (cognitive) tasks, many individuals who have recovered from hearing loss (HL) perform at normal hearing (NH) levels while a subset remains impaired. This suggests that brain regions down-stream of the perceptual processing centers of the primary auditory neuraxis are sensitive to developmental HL. One such region, which is highly involved in language development, is the auditory striatum. We have recently reported that animals that have recovered from HL continue to show significant changes in cellular properties in the cortex and striatum. Thus we asked how these changes to cellular and synaptic properties affect the learning of an auditory task. We used the gerbil and a brain slice preparation to assess how changes to synaptic excitation and inhibition correlate with behavioral performance in animals as they learn to discriminate two amplitude-modulated tones. We found that the synaptic changes that occur in NH and HL recovered animals are extremely polarized but have their strengths altered in such a way that permits learning in both groups of animals. In fact, there appears to be a range of excitatory and inhibitory synaptic strengths that is correlated with an increase in the probability of eLTP expression. Learning occurs rapidly over these days, suggesting that increased eLTP expression in the striatum plays a role in the acquisition of this discrimination task. Delays to the learning-induced progression towards this range could account for the shallower learning curves observed in both NH and HL animals. The precocious onset of the progression towards this range could like-wise account for the steeper learning curves observed in animals with faster task acquisition. Thus these results demonstrate a plasticity mechanism by which changes to the excitatory and inhibitory synaptic set points in both NH and HL individuals move towards a range of synaptic function that promotes learning. The broad implications of this research are that it will demonstrate how HL-induced changes to corticostriatal plasticity leads to persistent cognitive-behavioral impairments during associative learning in some individuals and not others.

32. Neural activities in marmoset premotor cortex for vocal production during social communication

Lingyun Zhao, Xiaoqin Wang

The neural mechanism of vocal production has been a long-standing question in non-human primates. While most of previous studies have focused on subcortical regions and the limbic pathway, recent experiments have revealed evidence for the involvement of the frontal cortex during vocal production. Using marmoset monkeys as a model system, studies in our lab and others have found that neurons in the premotor cortex were modulated when marmosets produced phee calls (a long-distance contact call) in vocal exchanges when isolated from their social groups. It has been known that marmosets maintain frequent vocal interactions in rich social contexts using four major types of calls: phee, trill, twitter and trillphee. Given previous findings for phee calls, it is an important question whether neural activities in premotor cortex are correlated with the vocal production of other call types used in social communication by marmosets. Here we investigated this question by recording neural activities from free-roaming marmosets in the colony when the subjects maintained visual contact and engaged in natural vocal interactions with other individuals. We developed a wireless multi-channel neural recording system to enable reliable recording in the marmoset cages and used a targeted acoustic recording setup to isolate subject vocalizations from the background. We found local field potentials and single unit activities in premotor cortex modulated by each type of social communication calls. Some recording sites and a subset of units showed difference in modulation when marmosets were producing different types of calls, suggesting a representation of call type in the premotor cortex. Our results provided further evidence for the role of the frontal cortex in vocal production and communication.

33. Predicting the neural responses to speech in human auditory cortex using deep neural network models

Menoua Keshishian, Hassan Akbari, Bahar Khalighinejad, Jose Herrero, Ashesh D Mehta, Nima Mesgarani

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 processes in the brain. Recent advancements in machine learning and computational power have allowed us to utilize deep learning methods in a large variety of tasks. We investigated the utility of deep neural network models to predict neural responses to speech in human auditory cortex, with the goal of analyzing the learned networks to gain insight into the nonlinear mechanisms of the brain. The neural responses were recorded from the perisylvian auditory cortex of five patients undergoing surgery for the treatment of epilepsy, as they listened to continuous speech. As deep convolutional neural networks (CNN) have shown great promise in capturing non-linear relationships, we trained a CNN with rectified linear 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.

In comparison to the STRFs, the predicted responses from the neural networks had a higher correlation with the original responses. On average, using CNNs improved performance by 25%. We analyzed the nonlinear function that the network implements to determine the computation of the auditory pathway and identified several properties that differentiate the linear and nonlinear functions. In addition, we examined the response dimensions in which the DNN excels at prediction, compared to the linear model. Finally, we examined the relation between the properties of the linear and nonlinear functions and the electrode anatomical locations. This study further shows that by interpreting complex nonlinear models that outperform their linear counterparts, we can gain invaluable insight into the workings of the brain, that are simply not possible with the traditional linear approach.

34. Population responses underlying statistical inference in a map of auditory space in the barn owl's midbrain

Roland Ferger, Michael V. Beckert, Keanu Shadron, Daniel Sanculi, William M. Debello, Brian J. Fischer, Jose L. Pena

The barn owl is a nocturnal hunter with outstanding sound localization abilities. It has become a model organism to study neural circuits of extracting binaural cues for sound localization that are relevant for behavior. Neurons in the owl's midbrain are known to respond maximally to acoustic stimuli with a distinct combination of interaural time difference (ITD) and interaural level difference (ILD). Together, these neurons form a neural map of auditory space which supports sound-orienting behavior. However, open questions regarding how the neural population in the map is read out on a trial-by-trial basis remain unanswered. This work reaches beyond responses of single neurons to understand the relationship between the activity pattern across the neural population and behavior.

Recent work has shown that the read out of population activity by a population vector (PV) predicts orienting head saccades, and approximates Bayesian statistical inference by integrating the overrepresentation of frontal directions and the differential shape of spatial tuning curves across the map. We investigated whether the trial-by-trial variability of neural activity matched predictions made under a Bayesian model, which matches a PV to the behavioral output. When a stimulus becomes less reliable, for instance if ITD detection is disrupted by adding independent noise at the two ears, an animal's performance becomes less accurate. The Bayesian model predicts that this should manifest in a broadening of the population response in the map. An alternative hypothesis that could explain this behavioral effect is that the reduced reliability induces shifts in the population response, which are indistinguishable from a response to a different ITD, termed differential correlations. Using a microelectrode array to record multiple units across the map, we tested whether the trial-by-trial population activity better matched the Bayesian model or the presence of differential correlations. On single trials the activity matched the pattern of activity predicted by the Bayesian model implemented by a PV. This was consistent across a number of stimulus conditions. Furthermore, the correlation structure of neurons in the map did not support a trial-by-trial shift in the center-of-mass as predicted for differential correlations. This provides additional support for the PV model.

35. Neural competition in auditory decoding of own- and other-vocalizations

Joris Dietziker, Matthias Staib, Sascha Frühholz

Listening to the voices of other individuals (other-vocalizations) was reported to elicit increased activity in different neural systems, such as the auditory cortex (AC). Reduced activity in the AC was shown for speakers listening to their own voice while producing vocal utterances (own-vocalizations). Listening to own- and other-vocalizations while speaking could therefore lead to contrary effects on the neural auditory system. Yet little is known about the neural dynamics of this auditory cortical competition between processing own- and other vocalizations. In this experiment we tested how this competition is resolved in the AC.

We conducted an fMRI study using human volunteers to investigate brain responses to own- and other-vocalizations being presented, first, in an active speaking task and, second, in a passive listening task. For the active speaking task participants were instructed to vocalize simple vowels that were fed back in real time to one ear (either left or right) while simultaneously being presented with a vocalization of another person to the same (same ear condition) or the other ear (other ear condition). In the passive listening task participants were presented with voice recordings of their own voice and the voice of another person again presented on the same or on separate ears.

Our findings did not show that activations by other voices in the auditory cortex are significantly reduced by actively speaking, as overall cortex activation was higher in the active task than the passive task and similar activations were found in the active task in primary and higher auditory areas for own and other vocalizations, regardless of their lateralization. There was, however, a linear decrease in the passive task when looking at the overall activation in AC in trials where two voices were presented to the same ear. When comparing active and passive tasks, Heschl's gyrus was shown to be persistently more activated during active trials.

The findings from this study possibly indicate that cortical competition by own and other vocalizations is not resolved by suppressing the speaker's voice, nor by a simple additive process. Instead, in the active condition the own voice seems to be accounted for before neural competition can occur, which could be explained by a precise forward model modulating the auditory cortical response to self-generated speech.

36. A neural ensemble correlation code for sound category identification

Monty Escabi, Mina Sadeghi, Xiu Zhai, Ian Stevenson

Humans and other animals effortlessly identify sounds and categorize them into behaviorally relevant categories, yet, the acoustic features and neural transformations that enable the formation of perceptual categories are largely unknown. Here we demonstrate that the correlations between neuron ensembles in the auditory midbrain (inferior colliculus) of unanesthetized rabbits reflect correlated structure in sound envelopes and that these statistics can contribute to the discrimination of sound categories. Five natural texture sounds (running water, bird chorus, crackling fire, rattling snake and crowd noise) were delivered over calibrated headphones and multi-channel neural recordings (16 and 32 channels) were obtained from the inferior colliculus. We first demonstrate that neuron ensemble correlations are highly structured in both time and frequency and can be decoded to distinguish sounds. A neural classifier was developed that uses neural correlations statistics between frequency order sites as the principal response features. The sound identification performance of the neural classifier improved with the sound duration with an evidence accumulation time constant of ~ 1 second and accuracy rates approaching 90%. Next, we develop a probabilistic time-varying framework for measuring the nonstationary spectro-temporal correlation statistics between frequency organized channels in an auditory model. In a 13-category sound identification task, classification accuracy is consistently high ($>80\%$), improving with sound duration and plateauing at $\sim 1-3$ seconds, mirroring human performance trends. The nonstationary short-term correlation statistics were more informative about the sound category than the time-average correlation statistics (84% vs. 73% accuracy). When tested independently, the spectral and temporal correlations between the model outputs achieved a similar level of performance and appear to contribute equally. These results outline a plausible neural code in which correlation statistics between neuron ensembles of different frequencies can be read-out to identify and distinguish acoustic categories (supported by NIDCD R01DC015138).

37. Electrocorticographic Responses to Vowel Sequences in Awake and Anesthetized States

Kirill Nourski, Mitchell Steinschneider, Ariane E. Rhone, Hiroto Kawasaki, Matthew A. Howard, Matthew I. Banks

Elucidating neural signatures of sensory processing across awareness states is a major focus in neuroscience. Clinically relevant conditions of altered awareness include sedation, loss of consciousness (LOC) under general anesthesia, natural sleep, and disorders of consciousness. Non-invasive studies in humans using the general anesthetic propofol to induce sedation and LOC have shown differential effects on auditory cortical activity. Propofol administration led to a greater reduction of activity in non-primary auditory cortex and auditory-related areas compared to primary auditory cortex in Heschl's gyrus. High spatiotemporal resolution of electrocorticography (ECoG) can extend results of non-invasive studies to delineate hierarchical organization of human auditory cortex. Previous ECoG work showed regional heterogeneity in the effects of general anesthesia on responses to click trains (Nourski et al., *NeuroImage* 2017, 152:78-93). In the present study, more complex and ecologically salient stimuli (vowel sequences) were presented in an active target detection task during administration of propofol. This study sought to further characterize auditory cortical responses as subjects transitioned from awake to sedated to unconscious state.

Subjects were adult neurosurgical patients with intracranial electrodes placed to identify epileptic foci. Data were collected prior to electrode removal surgery. Stimuli were sequences of five 100 ms vowels separated by 50 ms silent intervals presented during an awake baseline state and during propofol administration. Subjects were asked to press a button in response to occasional target stimuli. Depth of anesthesia was monitored using the Observer's Assessment of Awareness Scale and bispectral index. Regions of interest (ROIs) included core and non-core auditory, temporo-parietal auditory-related and prefrontal cortex. Activity was measured as averaged evoked potentials (AEPs) and high gamma (70-150 Hz) event-related band power. Envelope-following responses (EFRs) were measured as autocorrelation of the first-order differential of the averaged response at the 150 ms lag.

Vowels elicited AEPs throughout all ROIs in the awake state; high gamma activity was restricted to auditory cortex; EFRs were prominent in core auditory cortex. Sedation was characterized by a decrease in the number of responsive sites in auditory-related cortex and PFC, decrease in AEP amplitude in most ROIs and increase in AEP onset latency in non-core auditory and auditory-related cortex. LOC was characterized by a further decrease in the number of responsive sites in auditory-related cortex and PFC, decrease in AEP amplitude in most ROIs, increase in AEP onset latency in auditory cortex, decrease in the number of responsive sites and a decrease in EFRs in core auditory cortex. The extent of activation within auditory cortex in sedated and unconscious states was greater for the vowel sequences than reported previously for click train stimuli (cf. Nourski et al., 2017).

Overall sensitivity of cortical responses to propofol increased along the ascending cortical processing hierarchy. Marked attenuation of responses to sound in auditory-related cortex and PFC may represent a biomarker of LOC under general anesthesia. The findings refine and clarify results of previous non-invasive studies and serve as a foundation for examining changes in

sensory processing associated with general anesthesia induced by other agents, natural sleep and disorders of consciousness.

38. Understanding speech in background noise relies on similar processes to figure-ground segregation

Emma Holmes, Timothy D. Griffiths

Speech-in-noise (SIN) perception is a critical everyday task that varies widely across individuals and cannot be fully explained by the pure-tone audiogram, which measures thresholds for detecting sounds in quiet. This finding is unsurprising, because speech is usually heard at suprathreshold levels. Instead, the difficulty likely arises because listeners must separate speech from simultaneously-occurring background sounds. We predicted that the neural processes that underlie successful SIN might partially overlap with those that underlie successful auditory figure-ground segregation. As a first step, we examined how much common variance links speech-in-noise perception to auditory figure-ground perception, and how this relationship depends on the properties of the figure to be detected.

We presented sentences from the Oldenburg matrix corpus (e.g., "Alan has two old sofas") simultaneously with multi-talker babble noise. We adapted the target-to-masker ratio to determine the participant's threshold for reporting 50% of sentences correctly. Our figure-ground stimuli included one based on Teki et al. (2013; PMID 23898398) in which each 50 ms time window contains random frequency elements; participants detected elements (the "figure") that remained fixed for 300 ms. We also tested figures that changed in frequency over time, mimicking the formants of speech; participants had to discriminate gaps that occurred in the "figure" or "background" components.

Audiometric thresholds at 4-8 kHz accounted for 15% of the variance in SIN, yet figure-ground performance explained a significant portion of the variance in SIN that was unaccounted for by variability in audiometric thresholds. These results are consistent with the idea that common neural processes contribute to both speech-in-noise and figure-ground perception.

39. Investigation of pitch encoding neurons in the ferret auditory cortex

Quentin GAUCHER, Aleksandar Z. IVANOV, Mariangela PANNIELLO, Johannes C. DAHMEN,
Benjamin WILLMORE, Andrew J. KING, Kerry M. M. WALKER

Pitch is one of the most salient and behaviourally relevant perceptual features of sound. It is the foundation of musical melody, and it plays a key role in both human and animal communication. Previous research in ferrets has helped to elucidate how the pitch of artificial vocal calls is encoded by auditory cortical neurons. We have shown that ferrets can classify the pitch of sounds as “low” or “high” (Walker et al., 2009), that neurons which are sensitive to pitch cues are distributed widely across the auditory cortex (Bizley et al., 2009, 2010; Walker et al., 2011), and that auditory cortical neurons represent ferrets’ trial-to-trial pitch judgments (Bizley et al., 2013). However, these studies did not examine whether the “pitch-sensitive” neurons were able to maintain their tuning to a preferred fundamental frequency (F0) across a variety of stimuli (“pitch-selective” neurons), like those described in the marmoset pitch area (Bendor & Wang, 2005). In this study, we investigated whether pitch-selective neurons may also exist in non-primate auditory cortex. We recorded the responses of large populations of individual neurons to a variety of sounds that each varied in F0 (17 F0 values; 250 - 4000 Hz). The stimuli included click trains with varied temporal periodicity, pure tones, as well as sounds with filtering and phase manipulations to manipulate resolved harmonic and temporal envelope cues. Data were collected from ketamine and medetomidine anesthetized adult ferrets using high-channel-count multielectrodes (Neuropixels) and by imaging single neuron calcium dynamics with GCaMP6 under a 2-photon microscope. These experiments identified a subset of auditory cortical neurons with pitch selective responses, similar to those previously reported in marmoset (Bendor & Wang, 2005). Ferret pitch-selective neurons usually showed greater sensitivity to temporal cues than resolved harmonic cues, in keeping with our recent behavioural findings (Walker et al., submitted).

40. Mapping the human subcortical auditory system: Validation across histology, ex vivo MRI, and in vivo MRI

Kevin R. Sitek*, Omer Faruk Gulban*, Evan Calabrese, G. Allan Johnson, Satrajit S. Ghosh†, Federico De Martino†

The human subcortical auditory system remains understudied, primarily due to technical challenges in imaging the human brainstem in vivo, as well as to limited human-specific anatomical information. Advances in MR imaging, especially at high magnetic fields, have greatly improved our imaging of subcortical structures. Yet existing knowledge about the human subcortical auditory pathway, mostly originating from histology, is difficult for many researchers to apply to their MRI investigations.

Our first aim in this work is to provide an openly available 3-D atlas of the human subcortical auditory system based on three complementary types of data: histology, ex vivo anatomical MRI, and in vivo high-field (7 Tesla) functional MRI. Ex vivo MRI was collected at the Duke Center for In Vivo Microscopy, with anatomical images at 50 μm isotropic resolution. In vivo 7T functional MRI was collected at Maastricht University at 1.1 mm isotropic resolution. These MRI images were compared to 100 μm Big Brain histology data from the Human Brain Project. We find large agreement across the three imaging methods, suggesting that MRI is sensitive to anatomical contrasts delineating the auditory pathway, and that 7T fMRI in particular is capable of uniquely identifying subcortical auditory structures along the pathway. A publicly available 3-D atlas will help future investigations by focusing on auditory structures based on a priori anatomical definitions.

After confirming the anatomical location of subcortical auditory structures in ex vivo and in vivo MRI, we next sought to validate the connections between these structures using diffusion MRI tractography. Ex vivo diffusion images from Duke CIVM were acquired at 200 μm isotropic resolution in 120 directions with a gradient strength $b = 4000$. In vivo diffusion MRI from Maastricht was acquired at 1.05 mm iso. resolution with $b = 1000, 2000, 3000 \text{ s/mm}^2$, in 66 directions and 11 additional $b = 0$ for each b -value (whole brain coverage, ~ 60 min. acquisition time). In both ex vivo and in vivo MRI, we can reconstruct likely white matter connections between subcortical auditory structures.

Each of these datasets—histology, ex vivo MRI, and in vivo MRI—provides valuable, unique information about the structure and connectivity of the subcortical auditory system. Their agreement suggests that in vivo MRI is sufficiently capable of identifying auditory nuclei. We expect that our atlas will be very beneficial to new investigations into the function and connectivity of the human subcortical auditory system.

41. Neurons in auditory cortex are sensitive to frequency pattern violation

Linda Garami, Chris Angeloni, Kathrine C. Wood, Maria N. Geffen

The human auditory system efficiently integrates spectral and temporal features of the environment, and auditory objects are often detected through regularities in their acoustical structure. Yet the underlying neural mechanisms remain unknown. Our goal was to identify the neuronal circuitry in auditory cortex (AC) that enables the auditory system to detect regularities in sounds.

It has been shown that a network of auditory and frontal brain areas detects regular acoustic patterns. At a gross level, imaging studies showed that the AC and the prefrontal cortical regions exhibit sensitivity to violation of spectro-temporal acoustic regularities. Responses of individual neurons in AC also exhibit sensitivity to temporal regularities in sounds: Individual neurons exhibit stimulus-specific adaptation (SSA), a reduction in their response selective to frequently presented inputs. This adaptation may underlie the population sensitivity to more complex spectro-temporal acoustic regularities. Interneurons, such as Parvalbumin- or Somatostatin-positive neurons, through differential post-synaptic integration, can amplify adaptation to spectro-temporal patterns in excitatory neurons.

Here, we adapted an oddball paradigm used for testing prediction errors in humans, to test the hypothesis that neurons in AC detect and encode pitch pattern violations. We used an oddball set up for probing the encoding of pitch structure. Tone pip frequencies of F1-4 were selected at 0.2 octave intervals, narrower than the typical tuning bandwidth of neurons in AC. The 4 different tone pips were arranged in pairs in a way that F1 was either preceded by a tone with an intermediate (standard) or with a large (deviant) difference in pitch. We recorded neuronal activity in AC of head-fixed mice using an electrode array and compared the strength of responses to the deviant and standard to those in the equal probability condition.

In our analysis for F1, we identified a group of neurons across 4 mice that showed stimulus-specific adaptation for this frequency. A subset of these neurons showed a firing rate increase when F1 was preceded by a rare tone (large pitch difference) compared with a standard tone (intermediate difference), suggesting that AC neurons are sensitive to spectro-temporal patterns and that the firing rate is modulated by the violation of the statistical pattern based on the frequency. These results advance our understanding of how an auditory figure would emerge based on adaptation to the spectro-temporal regularities.

42. Effects of peristimulus vagal nerve stimulation on responses in ferret auditory cortex

Jonathan B. Fritz, Ali Mohammed, Jaya Viswanathan, Pingbo Yin, Diego Elgueda, Erin Causey, Jeslyn Lai, Stephen V. David, Shihab Shamma

Vagal nerve stimulation (VNS) has been shown to be an effective therapy for treatment of anxiety, epilepsy, inflammation and improves rehabilitation following stroke or brain injury. VNS is known to activate the Nucleus Tractus Solitarius which elicits CNS release of the neuromodulators acetylcholine (Nucleus Basalis) and norepinephrine (Locus Coeruleus), each of which are known from previous studies to affect responses in auditory cortex and enhance neuroplasticity. The goal of the present research was to explore the cortical effects of VNS in the awake animal, which could provide the neurobiological basis for VNS targeted neuroplasticity training – the use of VNS to enhance auditory learning. In the present study, we investigated the effects of VNS on auditory cortical responses in the awake, quiescent ferret to tonal, noisy and Mandarin Chinese phonemic stimuli with or without pairing with peristimulus VNS. We varied VNS duration, interstimulus interval, current amplitude and stimulation rate to explore parameter space for optimal stimulation effects. We also varied the site of VNS stimulation, using either cuff stimulation of the cervical vagus (c-VNS) or transdermal stimulation of the concha of the external ear, known to be innervated by the auricular branch of the vagus (a-VNS). We recorded from 120 neurons in primary auditory cortex (A1) and found that 32/120 cells showed response enhancement of 30% or more and 12/120 showed suppression effects of 30% or more. In some neurons, VNS lead to striking enhancement of cortical responses by 100-200%. Effects lasted for minutes and sometimes for up to one hour post-VNS, but gradually decreased and cell responses returned to baseline. Effects could be replicated with a second application of VNS. A parallel set of effects was observed with pupillary responses which were also modulated by VNS. In addition to measuring effects of VNS in a quiescent, listening animal, we also measured responses to VNS during active auditory task engagement in auditory go-nogo tasks, pairing target stimuli with or without peristimulus VNS. We shall describe the differences between the effects of a-VNS and c-VNS on cortical responses to acoustic stimuli during active behavior and passive listening.

43. A computational model of the underlying mechanisms of temporal coding in auditory cortex

Jong Hoon LEE, Xiaoqin Wang, Daniel Bendor

Neurons in the mammalian primary auditory cortex can temporally lock to individual acoustic events up to around 40-50Hz, covering the range of acoustic flutter. While modulation rates within the perceptual range of acoustic flutter are therefore represented temporally, primary auditory cortex neurons can also monotonically increase (Sync+) or decrease (Sync-) their discharge rates over this range of repetition rates that span flutter perception (Bendor and Wang 2007). Although neural mechanisms of stimulus-synchronized and non-synchronized responses have been studied using computational models of single neurons (Bendor 2015, Gao and Wang 2016), the integration of rate coding in stimulus-synchronized responses, to generate Sync+ and Sync- responses, has not yet been directly examined using such computational models. Here we investigated the underlying neural mechanisms responsible for Sync+ and Sync- responses in auditory cortex, and demonstrated that the addition of synaptic depression to a leaky integrate-and-fire excitation-inhibition model can reproduce these two response modes. Specifically, stronger synaptic depression of excitatory inputs relative to inhibitory inputs leads to Sync- responses while weaker synaptic depression of excitatory inputs relative to inhibitory inputs leads to Sync+ responses.

44. Attentional modulation of neural speech tracking and alpha power independently support speech comprehension in middle-aged adults

Sarah Tune, Mohsen Alavash, Lorenz Fiedler, Jonas Obleser

Speech comprehension in noisy, multi-talker situations is facilitated by the engagement of auditory selective attention and by the generation of context-based semantic predictions. Here, we investigated how healthy aging and its associated gradual decline in sensory acuity and cognitive functioning affect the relative balance of these cognitive strategies. Specifically, we asked whether age changes at the performance level would be mediated by concomitant changes at the neural level.

In an EEG study focused on healthy middle-aged human adults (N=73; 39–70y), participants listened to two competing, dichotically presented sentences. They were probed on the last word in one of the two sentences. Critically, two visual cues preceded auditory presentation. First, a spatial-attention cue either indicated the to-be-probed side, thus invoking selective attention, or 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 final word of both sentences, thus allowing to utilize a semantic prediction.

At the behavioral level, we observed a general performance benefit for informative compared to uninformative cues. Following a specific semantic cue, participants responded faster but not more accurately. Selective-attention cues lead to more accurate and faster responses. In addition, this relative, cue-related benefit in performance increased with age.

The analysis of EEG data revealed a similar modulation of neural dynamics by the two cues. Selective attention trials yielded a pronounced lateralization of 8–12 Hz alpha power during sentence presentation. Moreover, in selective attention trials the to-be-attended sentence envelopes were neurally tracked more closely, as indexed by better reconstruction accuracies in a linear decoding model. However, reconstruction accuracy was co-modulated by the semantic cue: under divided but not selective attention, reconstruction was more accurate for the specific compared to the general cue. Notably, both the degree of alpha lateralization and neural speech tracking predicted task performance but they did so independently of one another and independent of age.

In sum, our results provide evidence for the joint influence of attentional control and semantic predictions on listening success. They point towards a segregation of age-related changes at the neural and performance level in middle-aged adults. Moreover, our results indicate an independent attentional modulation of low-frequency auditory tracking and the lateralization of alpha power.

45. Interactions between intra- and inter-areal connections of human insula in processing emotional sounds

Yang Zhang, Wenjing Zhou, Juan Huang, Bo Hong, Xiaoqin Wang

Previous studies have demonstrated that the human insula are involved in two complementary sub-networks (Zhang et al., 2018 and Cauda et al., 2011), in which the posterior insula (PI) is connected to the sensory cortex and shows selective responses to sensory stimuli including auditory stimuli, where the anterior insula (AI) is connected to the limbic system and shows selective responses to emotional stimuli. To dissect the contributions of PI-to-Heschl's gyrus (HG) and AI-to-Amygdala sub-networks during emotional and non-emotional sounds processing in passive and active attention conditions, we applied conditional Granger causality analysis to assess the directional influences among the neural signals (iEEG) simultaneously recorded from HG, PI, AI and Amygdala of epilepsy patients. Our results showed that connections of the PI-HG sub-network were significantly stronger for emotional stimuli than non-emotional stimuli, but attention level showed no influence. While the connections of the AI-Amygdala sub-network were significantly influenced by both stimuli types and attention level. Furthermore, we found the connections of the PI-HG sub-network were highly dependent on the feedback signals from the AI-Amygdala sub-network during emotional stimuli condition, where the PI-HG sub-network had no influence on the connections of the AI-Amygdala sub-network. In addition, we found the connections between sub-networks are influenced by lateral interactions within the sub-network. Our findings suggest that feedback signals from higher order sub-network and lateral connections within the sub-network closely interact to mediate emotion perception from auditory stimuli.

46. Alpha oscillations index the temporal dynamics of exerted cognitive effort during listening

Björn Herrmann, Burkhard Maess, Ingrid S. Johnsrude

The ability to exert cognitive effort during listening is crucial for speech perception in the presence of background sound, and may be an important factor that determines listening success in older people with hearing difficulties. Neural alpha oscillations in parietal cortex may provide a promising index for the assessment of effort as its power increases when individuals listen for subtle acoustic changes in sounds. It is, however, less clear how alpha oscillations are modulated by prior knowledge about when in time cognitive effort must be exerted during listening. In electroencephalography (EEG) and magnetoencephalography (MEG) experiments ($N > 100$), we investigated how alpha power in a gap-detection task is affected by knowledge about when a near-threshold gap occurs within 10-s white noise sounds. Within one recording block, the gap occurred either within the first or the second half of the sound, and participants (21–33 years) were informed prior to each block. The precise position of the gap within the specified half was unknown to participants. Reaction times indicate that participants shifted their attention to either the first or second half of the sound, depending on the anticipated gap occurrence. EEG data showed a peak in alpha power at parietal electrodes either within the first or the last 5 seconds after sound onset, depending on whether participants were instructed to focus on the first versus second half of the sound. When detecting supra-threshold gaps, reaction times again indicated that participants shifted their attention to either the first or second half of the sound, but alpha power was less sensitive to this manipulation. These results suggest that investment of effort is needed to modulate alpha power. MEG data show that alpha power in parietal cortex was sensitive to the manipulation of gap occurrence (first vs. second half), but that alpha-power in auditory cortex remained enhanced throughout sound presentation (relative to baseline), independent of gap occurrence. MEG data from older people (54–72 years) show similar context-dependent modulations of brain activity, but in inferior parietal cortex. The data show that alpha oscillations in parietal cortex are sensitive to when in time cognitive effort is exerted during listening.

47. Non-reciprocal open-loop interactions in thalamo-thalamic reticular network

Kush Paul, Jeffrey W. Brown, Aynaz Taheri, Robert V. Kenyon, Tanya Y. Berger-Wolf, Daniel A. Llano

The thalamic reticular nucleus (TRN) has closed loop reciprocal connectivity with the thalamus which is well established. However, there is a growing evidence of additional non-reciprocal or open loop connectivity between them, in which the thalamic neurons are not reciprocally inhibited by the TRN neuron that they excite. Such an arrangement may provide a basis for the transmission of neuronal information between thalamic nuclei, both intramodal and crossmodal. In this study, we employ experimental and computational approaches to show that information transmission may occur along the thalamus via non-reciprocal open loop thalamus-TRN interconnectivity.

Whole-cell patch-clamp recordings were performed in colliculo-thalamocortical and somatosensory slice preparations 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 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.

Pharmacological investigations showed that monosynaptic and disynaptic TRN-thalamus inhibitory maps are attenuated by local application of excitatory synaptic blockers in the TRN but not by inhibitory blockers. This suggests that open-loop interactions occur via the TRN-thalamus AMPA connectivity. We further investigated contributions of TRN-TRN electrical synapses in the non-reciprocal transmission using gap junction blockers as well as in Connexin 36 KO mice.

The experimental findings were complemented with a computational model of the thalamo-reticular-cortical network to show that propagation of oscillatory activity was best supported in networks with strong closed and open loop reticulothalamic connectivity with minimal contribution of intrareticular chemical or electrical synapses.

48. A Cortical Time-Frequency Mask for Auditory Scene Analysis

Kenny F Chou, H Steven Colburn, Kamal Sen

Those who suffer from hearing-impairment lose the ability to focus on a single target in the presence of competing sounds. To address this issue, many computational auditory scene analysis (CASA) algorithms for sound segregation have been explored. The goal of many CASA algorithms is to estimate the “ideal binary mask”, which attenuates all non-target dominated time-frequency spectrogram tiles. These algorithms tackle the problem from an engineering perspective and have little biological relevance. In this work, we present a physiologically-based algorithm for separating spatially distributed sound sources via estimating a time-frequency mask. Our algorithm first encodes the sound mixture into spike trains, and separate its components according to their spatial locations, using a model of spatially tuned neurons in the avian midbrain. The spike train corresponding to the target of interest is then read-out by a cortical network model, based on the songbird auditory cortex analog. Finally, the output cortical spike trains are used to construct a time-frequency mask to isolate the target within a mixture. Our algorithm does not require training, retains binaural information, and can produce sounds with good intelligibility. This algorithm can potentially be used to improve the performance of hearing assistive devices, e.g., hearing aids and cochlear implants, as well as speech recognition technology, for auditory scene analysis.

49. Detecting tinnitus in nonhuman primates by using a non-acoustic startle paradigm

Lars Rogenmoser, Pawel Kusmierrek, Denis Archakov, Josef P. Rauschecker

Tinnitus impairs the quality of life of millions of Americans, making it a current concern for Public Health in

an aging population. Animal models are indispensable for the development of evidence-based therapy. Existing rodent models of tinnitus have been criticized because their results failed to translate to human patients. The reason for this may be that rodents lack a brain area in medial prefrontal cortex (vmPFC) that human imaging studies have shown to be causally related to the gating of tinnitus. Since vmPFC is highly developed in nonhuman primates, we aimed to establish a tinnitus model in rhesus monkeys. Tinnitus was determined by using a non-acoustic startle paradigm. In contrast to most animal studies in which the startle response is recorded by an accelerometer, we measured the eye blink, as commonly done in human startle experiments. Eye blinks were monitored by recording electromyographic (EMG) activity in response to air-puffs as startle stimuli, preceded by short auditory stimuli varying in frequency and intensity. The tones were adjusted according to the hearing thresholds, which were determined by frequency-specific Auditory Brainstem Response recordings. The threshold-adjusted

intensity levels were: +30dB SL, +6dB SL and -6dB SL. Since tinnitus loudness is known to range between

6-30 dB SL, we expected tinnitus to mask the mimicking frequency at the +6dB SL level, revealing the tinnitus frequency by an altered startle response. In this pilot study, one monkey was tested at its baseline, at a reversible tinnitus level (after administration of salicylate, 200mg/kg), and at a follow-up level. In order to ensure translation of the results to humans, a sample of human tinnitus patients and of matched control subjects without tinnitus underwent the same testing paradigm. The peaks of the EMG activity were extracted and subjected to inferential statistics. Unlike previous studies on tinnitus that make use of the Prepulse Inhibition phenomenon for tinnitus detection, our preliminary data strongly suggest the opposite, namely a Prepulse Facilitation. Our preliminary results suggest that the preceding tone facilitates the eye blink response as long as it is reliably perceived. In both species, the +6dB adjustment revealed the tinnitus frequency by a lack of facilitation.

The use of a non-acoustic startle stimulus is advantageous since it is free from acoustic interference and less aversive (especially for patients with hearing issues like tinnitus and hyperacusis). Since startle paradigms do not require instrumental conditioning or training, this set-up could easily be applied to a larger population, such as geriatric monkeys in a primate center.

50. Inhibitory inputs to medial geniculate body modulate nonlinear population auditory cortical responses after midbrain stimulation

Baher A. Ibrahim, Aynaz Taheri, Robert V. Kenyon, Tanya Berger-Wolf and Daniel A. Llano

The auditory colliculo-thalamocortical mouse brain slice, which retains synaptic connectivity between inferior colliculus (IC), medial geniculate body (MGB), and auditory cortex (AC), enables the examination of the role of the MGB in signal transmission from the IC to the AC. Low-magnification flavoprotein autofluorescence and calcium imaging as well as cortical local field potentials showed all-or-none cortical responses without any change in IC and MGB responses following IC stimulation. Low stimulating current amplitude and short inter-stimulus-interval increased the frequency of the missing cortical responses. In contrast, direct stimulation of MGB or the white matter produced only linear cortical responses. Bath perfusion of gabazine, GABA_A receptor blocker, was capable of retrieving the missing cortical responses. However, focal injection of gabazine into MGB, not AC, retrieved the missing cortical responses and linearized cortical and MGB responses. Further, current and voltage clamp recording from layer 4 or 2/3 showed similar all-or-none responses with showing no evidence for local inhibitory events during the missing cortical responses. However, voltage clamp recording of MGB cells showed that MGB receives a higher inhibition/excitation ratio during the missing cortical response. Interestingly, opto- and chemogenetic silencing of NTSR-positive layer 6 corticothalamic neurons retrieved the missing cortical responses which modulates feedback inhibition of MGB to control the gain of thalamocortical neurons. Further, current clamp recording of MGB cells showed that not all MGB cells fire in a response to IC stimulation. Similarly, calcium imaging of MGB showed a spatiotemporal heterogeneity of MGB cellular responses following each IC stimulation, and the latencies of calcium signal peak of MGB cells had a higher variance during the missing cortical responses. Network analysis of the MGB responses based on calcium imaging showed that the size of the network shrinks to intensify the edges between certain nodes during cortical activation only. These data suggest that the thalamus may recruit cortical ensembles rather than linearly encoding ascending stimuli, and that GABAergic cortical thalamic inhibition from the thalamic reticular nucleus may play a role in selecting cortical ensembles for activation.

51. Correlations improve accuracy in predicting population codes in macaque auditory cortex

M. B. Schaff, S. Subramanian, C. Lo, O. Contreras, E. Piasini, V. Balasubramanian, Y. E. Cohen

Our understanding of the function of the auditory cortex derives almost exclusively from experiments that study the coding properties of single neurons. Although these studies have shed important information, they are limited because they have not considered how neuron-to-neuron (i.e., pairwise) correlations may affect population coding. We, therefore, sought to determine whether, in the primate auditory cortex, pairwise correlations contribute information to population coding above and beyond the information provided by the independent firing of auditory-cortex neurons. We recorded neuronal activity in the primary auditory cortex and the anterolateral belt region of the auditory cortex in two rhesus monkeys. Neuronal activity was recorded either with tetrodes or multi-contact linear u-probe electrodes. While recording neuronal activity, the monkeys participated in a “hearing-in-noise” task in which they reported hearing a target stimulus that was embedded in a noisy background. We titrated task difficulty by varying the sound level of the target stimulus, relative to the background. Based on either (1) the independent spiking activity of these neurons or (2) the independent spiking activity plus their pairwise correlations, we generated “Ising” models. Ising models are the maximum-entropy distributions that are consistent with experimentally observed firing rates and pairwise correlations without any mechanistic assumptions. We report whether Ising models more accurately predict the experimental probability of population dynamics when pairwise correlations are taken into account along with independent spikes. We also test whether Ising models better represent population dynamics in primary auditory cortex than in the anterolateral belt. This study lays groundwork for future studies that determine the relative importance of correlations in processing different types of acoustic stimuli.

52. Active tracking of sound textures: a human intracranial study

Alexander J. Billig, Phillip E. Gander, William Sedley, Maria Chait, Hiroto Kawasaki, Christopher K. Kovach, Matthew A. Howard, Timothy D. Griffiths

A network consisting of auditory cortex, hippocampus, and inferior frontal gyrus has been proposed to support auditory working memory (Kumar et al., *J Neurosci* 36:4492-4505, 2016). However, human intracranial recordings are yet to reveal hippocampal high gamma activity, a neural spiking correlate, during memory for tones. Recordings from the rat hippocampal complex indicate that cells that support navigation can also form discrete firing fields for particular sound frequencies (Aronov et al., *Nature* 543:719-722, 2017). Critically, this is the case only when the animal adjusts a sound to match the frequency of a remembered target, but not during passive listening. For humans performing such a task, recruitment of navigation circuits may stem from a cognitive association between height in pitch and in physical space (Rusconi et al., *Cognition* 99:113-129, 2006). Here, we study the neural underpinnings of memory for and tracking of sound textures that vary continuously in a single dimension but have no such spatial association. Stimuli were concatenated chords, each containing between 4 and 100 simultaneous 200-ms tones randomly distributed in frequency over a 4-octave range. Fixing the number of simultaneous tones (“density”) while varying their frequencies from chord to chord gave rise to textures that were more “beep”-like at lower densities and more “noise”-like at higher densities. In each active trial, human subjects (patients implanted with electrodes to localize epileptic activity) listened to a 2-s sound of fixed target density, which they were to remember over a subsequent 2-s retention period. They then heard a 15-s texture, the density of which they adjusted using button presses to match the density of the target. Preliminary analyses of sound-induced neural activity revealed robust (high) gamma responses in Heschl’s gyrus and superior temporal gyrus, with concurrent low frequency (4-30 Hz) power decreases over a range of frontal, temporal, and parietal sites. (High) gamma responses were also observed in right hippocampus and right inferior frontal gyrus; these were greater during active adjustment than in a passive listening condition with similar auditory stimulation. Activity at one hippocampal site was also present during target presentation and the subsequent silent retention period. Ongoing work includes conducting motor and attention control experiments, and using multivariate analyses to study the link between behaviour and the strength of target representations across the putative auditory working memory brain network.

53. A potential role of non-lemniscal auditory thalamic inputs to lateral amygdala in diverse sound-induced defensive behaviors

Miaomiao Liu, Yiwei Wang, Hanqing Wang, Dongqin Cai, Ling You, Fenghua Xie, Yin Yue, Kexin Yuan

Amygdala is essential for formation and storage of the memory associated with fearful or rewarding environment stimuli. Lateral amygdala (LA) is a major input region of the amygdala. Decades of studies have shown the importance of LA in auditory fear conditioning, and it was proposed that the unconditioned-stimulus (US) and acoustic conditioned-stimulus (CS) converge in this region. Once the animal is conditioned, the acoustic CS, which is believed to be mainly relayed from the non-lemniscal auditory thalamus, triggers freezing behavior. In present study, we investigated whether the non-lemniscal auditory thalamus→LA pathway itself would be involved in defensive behaviors without being paired with US. We found that selective optogenetic activation of the non-lemniscal auditory thalamus→LA pathway led to avoidance and anxiety-like behavior, rather than freezing. Specifically, during real-time-place-avoidance test, the tested mice would run away from and mostly stay out of the laser-on chamber once they have explored that chamber for a few times. Therefore, they spent most of their time in the laser-off chamber. Control mice spent almost equal time in the two chambers. During open-field test, activation of this pathway increased the chances of rearing-up and reduced the entrance of central area. In contrast, control mice demonstrated normal exploring behavior. To our knowledge, this is the first study demonstrating independent involvement of non-lemniscal auditory thalamus→LA pathway in avoidance and anxiety-like behaviors. Our results suggested that, although the non-lemniscal auditory thalamus→LA pathway alone was not sufficient to mediate freezing behavior, this pathway might still play an important role in other sound-induced defensive behaviors.

54. Variant-Independent Sound Representation and Neural Discriminability in the Zebra Finch Auditory Forebrain

Mingwen Dong, Mimi Phan, David Vicario

Discerning the neural mechanisms behind variant-independent sound representation is essential for understanding spoken language and other forms of complex auditory communication. Zebra finches are the best-developed model system for studying the neural basis of processing communication signals; adult finches sing highly structured and unique learned songs that are used for individual recognition and mate selection. Although each individual's song is a set of sounds produced with a fixed pattern, actual songs contain certain variations, much like the same word produced by different speakers. Past studies have shown that neural responses in a higher auditory area (NCM) of the zebra finch adapt as novel stimuli become more familiar, suggesting that NCM plays an important role in encoding and remembering individual songs. We propose that NCM neurons exposed to variants of the same song lose the ability to discriminate those variants with repeated presentation and that this process leads to a variant-independent sound representation. To test this hypothesis, we recorded neurons from 32 electrodes placed bilaterally in NCM and Field L (a primary auditory region analogous to A1 that projects to NCM) during song presentation in awake, restrained birds. The birds heard either blocked repetitions of certain song variants (blocked), randomly ordered song variants (shuffled), or variants of 2 different songs in random order (contrast). After thresholding to extract multi-unit spikes and computing the response to each repetition of a given stimulus, we analyzed neural decoding accuracy using two related methods: one based on Euclidean distance between the response temporal profiles and a K-nearest neighbor machine-learning algorithm. Each algorithm's accuracy at predicting the stimulus that elicited a given neural response was used as a metric for neural discriminability. We found that across stimulus trials, NCM neurons in the shuffled and contrast condition gradually lost the ability to discriminate between variants of the same stimuli in a negative parabolic fashion while neurons exposed to the blocked condition exhibited less dramatic changes in terms of decoding accuracy.

55. Affective whispered voices in the human limbic system: differences in single-cell firing and LFP recordings from the amygdala and hippocampus

M. Bobin, M. Staib, T. Fedele, J. Sarnthein, S. Frühholz

Vocal communication is one of our most versatile and important ways of exchanging a broad variety of information. A speaker's affective state can be reliably decoded by a listener when perceiving the spectral properties of the speech. When perceiving affective vocalizations, a cerebral network including the auditory cortex (AC) and the medial temporal limbic (MTL) system classifies the different acoustic components of affective voices to match stored emotional pattern templates. However, it remains unclear how this neural dynamic identifies the actual emotion when the incoming acoustic signal is degraded. Listening to affective whispered voices is an acoustically sub-optimal situation for a clear comprehension, where the degraded vocal stimuli requires greater functional connectivity between the AC and the MTL. Activity in the MTL, especially when involving the hippocampus (HPC), should reflect an intensive retrieval from long-term memory information (e.g. past experiences, prototypes) in presence of affective voices reduced in acoustical quality. If so, providing this extra contextual memory association would require an acute engagement within the MTL of the bidirectional connected amygdala (AMY) and HPC. We recorded neural activity from nine drug-refractory epileptic patients, implanted with intracranial electrodes in the MTL for clinical purposes, recorded during an emotion identification task on normal and whispered affective voices. Neural spiking data and local field potential (LFP) activity corresponding to AMY or HPC channels were extracted and filtered. Spiking analyses revealed a higher proportion of HPC neurons firing following the onset of whispered voices, in comparison to voiced trials, while affective cues elicited significantly more AMY neuronal firing compared to the neutral condition. Furthermore, an increased coupling in the amplitude of LFP for the whispering condition was found between the AMY and HPC. For affective whispered voices, the neuronal firing from certain parts of the MTL was phase-locked to oscillations in the theta and the gamma bands. The combination from LFP and single-neuron spiking activity thus reflects an increased cooperation between the AMY and the HPC for decoding emotions from whispered voices, where the interpretation of an affective state is more demanding.

56. The frequency-band specific information flows in speech network

Yuxiang Yan, Hao Han, Dan Zhang, Wenjing Zhou, *Bo Hong

Characterizing the functional organization of speech network has been a focus of investigating the neural basis of language. Evidences have shown that there exists two cortical areas, roughly over superior temporal gyrus (STG) and inferior frontal gyrus (IFG), which are intensively involved during the speech related tasks, with STG more related to speech perception, and IFG to speech production (Hickok & Poeppel, 2007). Recent studies also found these two speech-related regions could be both activated in the perception and the production process (Cogan, 2014). However, the nature of interactions among this large-scale speech network, has not yet been elucidated. In this study, we recorded the intracranial EEG activities from 121 electrodes in brain-wide speech-related areas in three subjects, covering STG, IFG, anterior temporal lobe (AT), posterior temporal lobe (PT), parieto-temporal boundary (Spt) and motor areas. With high temporal resolution of the intracranial EEG, we characterized the dynamic activities in the speech networks, and quantified causal influences among these major network nodes during resting and speech-evoked states. In the intrinsic speech networks during resting, main information flows sourced from STG and IFG, and targeted to the other speech-related areas including AT, PT, Spt and motor areas. During speech perception and production tasks, strong information flow was found to modulate the high-gamma (60-140Hz) activities of STG & IFG ($p < 0.01$, compared to resting state), exhibiting a recurrent pattern of the functional connectivity between STG/IFG and other major nodes of the speech network. We also found that the causal connections in speech network showed a frequency-band specific feature: the causal connections in alpha - beta range (8-32Hz) were insensitive to task involvements, while the high-gamma band (60-140Hz) was strongly activated during the speech tasks ($p < 0.01$). In short, our data revealed distinct pattern of causal connectivity of speech networks during resting and speech task states, with separated frequency bands for networking.

57. Loss of interaural time difference sensitivity in rabbit inferior colliculus neurons following noise-induced hearing loss

Hariprakash Haragopal, Ryan Dorkoski, Gareth Whaley, Lukas Palmer, Mitchell Day

Sensorineural hearing loss (SNHL) is the most common form of permanent hearing loss, usually caused by cumulative overexposure to loud sounds. Individuals with SNHL show impairment in the abilities to both localize sound sources and use spatial cues to aid speech comprehension in noisy environments. However, the effects of SNHL on the circuits underlying sound localization are not known. Humans, and rabbits, localize sound in the horizontal plane (azimuth) via interaural time and level differences (ITDs and ILDs) of the sound waveform. These binaural cues are first encoded in the auditory brainstem, which projects to the inferior colliculus (IC). We measured responses of single-units in the IC of awake rabbits either with or without noise-induced hearing loss in response to auditory stimuli that varied in binaural cues in order to determine the effect of SNHL on sound localization coding. Rabbits were exposed to octave-band noise centered at 750 Hz at 133 dB SPL for 90 min under isoflurane anesthesia. Auditory brainstem response thresholds increased by approximately 30–45 dB between 0.5 and 16 kHz as measured two weeks after exposure, indicating extensive cochlear outer hair cell loss. Neurons from SNHL rabbits had greater sound-level thresholds to pure tones and noise stimuli than those from normal-hearing rabbits, as expected. Azimuth tuning curves were measured in response to noise stimuli with either natural ITD and ILD (ITD+ILD), or with one binaural cue fixed at zero (ITD-only and ILD-only). We quantified neural sensitivity to sound source azimuth (ITD+ILD stimulus) as the mutual information (MI) between firing rate and azimuth. There was a trend towards reduction in median MI for neurons from SNHL rabbits as compared to those from normal-hearing rabbits. We also computed MI between firing rate and ITD for noise stimuli that only varied in ITD (ILD fixed to zero). In this case, median MI was dramatically decreased for neurons from SNHL rabbits, with all MI values near zero, even though stimuli were above sound-level threshold. For neurons from SNHL rabbits, ITD+ILD azimuth tuning curves were very similar to ILD-only tuning curves, indicating neural sensitivity to azimuth was dominated by sensitivity to ILD. Our results suggest that 1) noise overexposure causes a loss of ITD sensitivity in the IC, and 2) any remaining information about sound source azimuth is due to sensitivity to ILD.

58. Closed-loop stimulation reveals the brain state-dependence of auditory perceptual sensitivity

Leonhard Waschke, Jonas Obleser

Instantaneous brain states, that is neural activity prior to stimulus onset, affect the processing of sensory information and subsequent conscious perception in a host of ways. Particularly, the degree of pre-stimulus neural synchronization as captured by low-frequency power is thought to affect the encoding and perception of auditory stimuli in non-human animal models, with less synchronization resulting in more thorough encoding. We recently demonstrated a strikingly similar pattern in human subjects by using time-resolved entropy of the EEG signal to approximate pre-stimulus neural synchronization and relate it to ensuing perceptual decisions (Waschke et al., Sci Rep 2017). In an endeavour to get closer to a mechanistic understanding of this state-dependence, tone stimuli were then presented into high- and low-entropy brain states, respectively, and subjects performed a challenging pitch discrimination task upon these tones.

We constructed a closed-loop experimental setup within which we recorded EEG and pupillometry data from human subjects (N = 25) and processed it using a near real-time algorithm: Time-varying entropy of EEG-signals from auditory cortical regions was calculated. A continuously adapting criterion was used to determine periods of relatively high (peaks) or low entropy (troughs). We expected that peak-entropy states would not only be accompanied by reduced low oscillatory power, but also lead to enhanced tone-evoked responses and neural encoding, and ultimately greater perceptual sensitivity (as determined by the modelled psychometric curve).

Replicating our previous result, we found higher entropy to be associated with decreased low-frequency oscillatory power and an increase in phase coherence shortly after tone onset. Furthermore, inverted u-shaped relationships between pre-stimulus entropy and both phase coherence and ERP magnitude were found, with medium entropy resulting in the strongest evoked response. Importantly, this pattern of neural activity was mimicked by behaviour: response speed was highest at medium levels of pre-stimulus entropy. The same relationship was present for perceptual sensitivity. In sum, pre-stimulus entropy likely impacts perceptual sensitivity by altering evoked activity and the encoding of sensory information.

These results validate a promising conceptual approach to quantify states of neural synchronization in the EEG and additionally offer new insights into the brain state-dependence of human perception. We discuss our findings in the context of rapid fluctuations in arousal and associated neuromodulatory changes.

59. Auditory texture synthesis from task-optimized convolutional neural networks

Jenelle Feather, Josh H. McDermott

Models of sensory systems have traditionally been built from engineering principles and experimental observations, but modern-day machine learning allows models to be learned from data. We sought to compare hand-engineered and learned models of the auditory system by generating synthetic sounds that were matched to particular natural sounds according to the features in the model, and used automated gradient-based optimization to synthesize sounds for this purpose. Models that replicate perceptual representations should produce model-matched sounds that sound like the natural sounds they are matched to. We synthesized stimuli that produce the same time-averaged values in a model's representation as those measured from a natural texture. Such stimuli should evoke the same texture percept if the model replicates the representations underlying texture perception. We compared textures generated from: (1) the first layer of filters from three task-optimized convolutional neural networks, (2) a model of primary auditory cortex consisting of spectrotemporal filters, and (3) the McDermott and Simoncelli (2011) texture model (consisting of cochlear and temporal modulation filters and their correlations). The task-optimized networks were trained on a word in noise task, a speaker identification task, or a music genre classification task. Sounds generated from any of the task-optimized filter banks were more recognizable and realistic than those from the hand-engineered filter bank. To explore the origins of this difference, we constrained sounds to additionally match marginal statistics of the cochlear representation. The inclusion of cochlear statistics caused the quality of sounds from the hand-engineered model to improve to the level of those from the learned filter bank. By contrast, including cochlear statistics did not improve the quality of sounds from the learned filters. Synthesis from the learned filters was comparable to that from the McDermott and Simoncelli model, which also required cochlear statistics to produce realistic textures. The trained filters evidently retain task-relevant information from earlier processing stages that is discarded by conventional models. Further, sounds generated from random filters were less recognizable and realistic than the task-optimized filters, suggesting that features learned during training are necessary for texture synthesis. The results illustrate that better auditory models can be obtained by task-optimizing sensory representations.

60. The neural processing of phonemes is shaped by linguistic analysis

Jackson C Lee, Tobias Overath

Speech perception entails the mapping of the acoustic waveform to stored linguistic representations, such as phonemes, syllables, or words. Recent evidence suggests that different phoneme classes (e.g. plosives, fricatives, etc.) have characteristic neural signatures, or phoneme-related potentials (PRPs; Khalighinejad et al., 2017). What remains to be understood is the extent to which the temporal scale of linguistic analysis, and linguistic knowledge, influence the processing of this fundamental linguistic unit.

To control the scale of linguistic analysis, we used a modification of our speech quilting algorithm (Overath et al., 2015) to generate stimuli that maintain linguistic structure at one of 4 linguistic units: phoneme, syllable, word or sentence; to control for linguistic knowledge, we constructed speech quilts from both familiar (English) and foreign (Korean) languages. We recorded EEG from 28 native English speakers (with no knowledge of Korean); data were epoched to the phoneme onset boundary. Grand average PRPs across all phonemes showed a similar sequence of P50, N100, and P200 components at fronto-central regions. Comparisons between linguistic-unit levels (phoneme, syllable, word, sentence), showed significant differences ($p < 0.05$) at time points corresponding to the P50 and N100 components for English, but not for Korean. In addition, the N100 component showed a main effect of language and an interaction between language and linguistic unit. The classification of phonemes based on articulatory manner revealed unique PRPs for each of these classes, and forms were similar between the two languages. The similarity in articulatory-class responses across languages suggests that acoustic features play an important role in the processing of phonemes. However, the main effect of language and the interaction with linguistic unit suggest that the processing of a fundamental linguistic unit, the phoneme, is already shaped by linguistic analysis as early as 100 ms after phoneme onset.

References:

- Khalighinejad et al. (2017), *J Neurosci* 37: 2176-2185.
Overath et al. (2015), *Nat Neurosci* 18: 903-911.

61. Decoding trajectories of multiple moving sound sources from EEG in cocktail party environment

Adam Bednar, Edmund C. Lalor

We can clearly perceive the location of moving auditory objects in space. However, the cortical representation of auditory space, and auditory motion in particular, is not well characterized. Recently, we showed that in an acoustic scene with one sound source, auditory cortex tracks the time-varying location of a continuously moving sound. Specifically, we found that both the delta phase and alpha power of electroencephalogram (EEG) can predict the sound source azimuth significantly above chance level (Bednar 2018).

However, in natural settings, we are almost exclusively presented with a mixture of multiple competing sounds and so we must focus our attention on the relevant source in order to segregate it from the background noise e.g. 'cocktail party effect'. While many studies have examined the neural underpinnings of attentional selection in a cocktail party problem - especially in the context of sound envelope tracking by the cortex, it is unclear how we process and utilize spatial information in complex acoustic scenes with multiple sound sources. In this study we aimed to answer two questions: (1) Can we decode time-varying locations of multiple concurrently presented moving sound sources from EEG? (2) How does selective attention influence the sound trajectory tracking in cortex?

Subjects listened to two simultaneously presented noise stimuli having different spectral content over headphones. The stimuli were acoustically modified to be perceived as randomly moving on a semi-circular trajectory in the horizontal plane. Participants were asked to pay attention to one of the two presented stimuli and detect embedded targets. 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.

Preliminary results show that locations of both attended and unattended sound sources are cortically represented and we can decode the trajectories of both sound stimuli from EEG with a reconstruction accuracy significantly above chance level. We also found that selective attention significantly enhances the cortical representation of the sound trajectory of the attended stimuli. In addition, we found correlation between our EEG-based trajectory reconstruction accuracy and behavior performance in the target detection task. We believe this method, extracting spatial information of the acoustic scene, has potential to complement established sound envelope tracking-based attentional decoders in cocktail party environments.

62. Neural voice decoding in the primate auditory cortex is task-dependent

Sascha Frühholz, Plamina Dimanova

The mammalian and especially the primate brain has an evolved auditory cortical system for voice processing which is located in the superior temporal cortex (STC) and referred to as “temporal voice area” (TVA). The TVA was so far found in human and non-human primates as well as in dogs. The TVA shows higher activity to vocal compared to other nonvocal sounds and has been assumed to be a cortical area with consistent and uniform voice-sensitivity across species. However, given its large cortical extension covering many areas in the primary, secondary, and high-level auditory cortex, we hypothesized that TVA is composed of functional subareas that decode voice information dependent on the current task requirements rather than representing a uniform and state-independent functional area. We, therefore, used functional magnetic resonance imaging while recording brain response in human volunteers that performed a complex cognitive (i.e. sound classification) or a simple acoustic task (i.e. intensity judgments) while listening to vocal and nonvocal sounds. We hypothesized that different tasks while listening to vocal sounds will elicit differential activity and connectivity of high-level and low-level auditory regions. We found that simple acoustic decisions on vocal compared to nonvocal sounds elicited higher activity in region Te1.0 of the primary auditory cortex (AC), while the complex sounds classification task elicited the highest activity in region Te3 of the higher-level AC. Activity in these regions correlated with the decision drift rate estimated with a standard drift-diffusion model on the behavioral data. Furthermore, while we found functional connectivity between bilateral auditory cortices for the simple decision task, especially left Te3 showed functional connectivity to bilateral inferior frontal cortices during the complex decision task. Taken together, rather than being a uniform functional area the present data highlight the notion that the TVA is a multi-functional and state-dependent cortical area including functional subareas and connectivity patterns that decode important voice and sound information that is most relevant for the current task.

63. Human Sound Localization Depends on Sound Intensity: Implications for Sensory Coding

Antje Ihlefeld, Nima Alamatsaz, Robert Shapley

A fundamental question of human perception is how we perceive target locations in space. Through our eyes and skin, the activation patterns of sensory organs provide rich spatial cues. However, for other sensory dimensions, including sound localization and visual depth perception, spatial locations must be computed by the brain. For instance, interaural time differences (ITDs) of the sounds reaching the ears allow listeners to localize sound in the horizontal plane. Our experiments tested two prevalent theories on how ITDs affect human sound localization: 1) the labelled-line model, encoding space through tuned representations of spatial location; versus 2) the hemispheric-difference model, representing space through spike-rate distances relative to a perceptual anchor. Unlike the labelled-line model, the hemispheric-difference model predicts that with decreasing intensity, sound localization should collapse toward midline reference, and this is what we observed behaviorally. These findings cast doubt on models of human sound localization that rely on a spatially tuned map. Moreover, analogous experimental results in vision indicate that perceived depth depends upon the contrast of the target. Based on our findings, we propose that the brain uses a canonical computation of location across sensory modalities: perceived location is encoded through population spike rate relative to baseline.

64. Two-Talker Attention Decoding from EEG with Nonlinear Neural Networks and Linear Methods

Gregory Ciccarelli, Michael Nolan, Joey Perricone, Paul Calamia, James O'Sullivan, Nima Mesgarani, Thomas Quatieri, Christopher Smalt

A particular challenge with hearing impairment is difficulty in attending to a single speaker in a multi-talker environment. This study is a confirmatory study that hypothesizes that nonlinear neural networks (NN) may provide an advantage to linear methods in auditory attention decoding from recorded electroencephalograph (EEG) waveforms.

Fourteen subjects gave written, informed consent to participate in an EEG-auditory collection protocol approved by the MIT review board. Each subject wore a 64 channel EEG cap and was presented with speech from two collocated talkers (one male, one female) reading different passages. Subjects were directed to attend to one of the two speakers. The audio presentation was randomly stopped and subjects were asked to repeat back the last sentence from the target speaker to check for adherence to the task. There were approximately 35 segments of approximately 45 seconds duration per subject.

We trained one neural network architecture after de Taillez (2017) to map from the collected EEG to the perfectly separated speech envelope, and we compared against a baseline linear method (O'Sullivan 2015). Attention was decoded by comparing the predicted auditory envelope using all 64 EEG leads to the target auditory envelope and the distractor envelope using Pearson correlation. A sliding window of 250 ms of EEG was used for decoding, and the subsequent predicted and measured envelopes were correlated using a 10 second window stepped in 10 second intervals to avoid overlap. Accuracy, as defined by whether the system correctly predicted the person was attending to the target talker, was determined for each 10 second interval. Training and testing was performed with leave-one-segment out cross validation.

We have analyzed fourteen subjects. Neural network decoding performance is significantly above chance for ten of the fourteen subjects and outperforms the linear least squares baseline for nine of these subjects. One subject was decoded with an 82.17% accuracy (NN) vs 75.54% (Linear baseline).

This work provides additional evidence that auditory attention decoding is possible from EEG, and additional evidence that a neural network architecture may provide better decoding accuracy than a linear baseline architecture.

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.

This material is based upon work supported by the Assistant Secretary of Defense for Research and Engineering under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Assistant Secretary of Defense for Research and Engineering.

65. Response to own name in noise differs in autistic adolescents with severe language impairments

Sophie Schwartz, Le Wang, Barbara Shinn-Cunningham, Helen Tager-Flusberg

Your name is one of the most common things you hear, beginning at birth. It becomes a highly salient stimulus, eliciting an orienting response that leads you to turn quickly towards its source upon hearing it. Children and adults show a positive event-related potential response (ERP) to their own name, but not other peoples' names, over parietal-occipital channels around 300 milliseconds post-stimulus (a P3), in attentive and inattentive states, and even in sleep and comatose states. However, recent evidence suggests that adults with autism (ASD) have a smaller P3 to their own name when compared to typically developing (TD) adults (Nijhof et al., 2018). These findings align with research showing that failure to orient to one's own name is an early sign of the disorder. Prior work has also found a relationship between the ability to segregate streams and extract salient speech from noise and integrity of language in ASD (Boatman et al., 2001; Russo et al., 2009; Wang et al., 2016). Our work expanded on prior findings by investigating ERPs to names in both quiet and multispeaker noise contexts, in TD listeners and ASD listeners with a range of language abilities. Subjects ages 13-23 heard 486 audio-recordings of their own and two other subjects' names in a passive ERP paradigm while they watched a silent movie. Additionally, we quantified expressive and receptive language in ASD subjects with parent report measures, including the Vineland Adaptive Behavior Scales and an in-house validated list of words used and understood (Plesa Skwerer et al., 2016). In quiet, TD subjects showed a P3 to their own name relative to other names during the first 30 presentations of each name ($p=0.05$), replicating prior studies that only presented this many trials. However, ERP topography changed as TD listeners were exposed to more repetitions of each name. From an analysis all usable trials, TD subjects showed a habituated frontal ERP to their own name at around 250 ms relative to other names ($p<0.05$) but no clear P3. When names were presented in noise, response to own name was significantly larger than to other names, lasting from 300-700 ms, along parietal-occipital channels (a prolonged P3) ($p<0.05$). In contrast, subjects with ASD showed no difference in response to their own and other names in either quiet or noise. Within ASD, receptive and expressive language positively correlated with P3 response to own name, but not other names, when embedded in noise ($p<0.05$). Future work should continue to explore how ASD individuals with co-existing language impairments process sound and if their ability to extract salient speech from noisy environments has any influence on their ability to process language.

66. Supra-threshold psychoacoustics and envelope-following response relations: Disentangling the contributions from outer-hair-cell loss and synaptopathy

Sarah Verhulst, Viacheslav Vasilkov, Markus Garrett, Frauke Ernst

The perceptual consequences of cochlear synaptopathy are presently not well understood as a direct quantification of synaptopathy is not possible in humans. To circumvent this problem and to study its role for human hearing, recent studies have correlated individual differences in subcortical EEG responses, as a proxy measure for synaptopathy, to changes in basic supra-threshold psychoacoustic tasks. However, it is not clear whether the reported missing relationships between EEG and psychoacoustic quantities are due to the adopted methods, or to a minor role of synaptopathy for sound perception.

To address this issue, we investigated the theoretical relationship between subcortical EEG and psychoacoustic methods for different sensorineural hearing deficits. We use a computational model of the human auditory periphery calibrated for auditory brainstem (ABR) and envelope following responses (EFRs) and simulate how different sensorineural hearing deficits impact EFRs and psychoacoustic cues important for amplitude-modulation (AM) and tone-in-noise (TiN) detection. Simulated synaptopathy can reduce EFR magnitudes and detection task performance by up to 15 dB, corroborating expectations from animal studies of synaptopathy. OHC loss was seen to impact the EFRs and psychoacoustic tasks to a much smaller degree, as these measures mostly rely on robust encoding of supra-threshold envelope cues. In fact, OHC deficits had a negligible role when OHC deficits and synaptopathy co-occurred. Our simulations show that individual differences in the considered metrics are explained by individual differences in the synaptopathy degree, irrespective of whether OHC deficits were also present. This finding is further supported by synaptopathy driving the regression between the EFR magnitude and AM/TiN detection thresholds across models with mixtures of OHC loss and synaptopathy.

These model-based conclusions are controversial, as several experimental studies question the role of synaptopathy for sound perception, but we present supporting data from human subjects with normal or sloping high-frequency audiograms. The individual differences in the EEG and psychoacoustic metrics match our predictions, and show that older normal-hearing listeners indeed perform as badly as the listeners with sloping audiograms, consistent with the view that both these groups suffer from synaptopathy.

67. Neural oscillatory mechanisms for interpersonal entrainment in music ensembles

Andrew Chang, Philip Chrapka, Dan J. Bosnyak, Laurel J. Trainor

People need to process complex dynamic information when coordinating with others in joint action, conversation and music. Previous EEG studies suggest that neural oscillations entrain to ongoing rhythmic temporal regularities in sensory input and predictively facilitate upcoming sensory perception and motor responses. However, existing studies have several limitations. First, most studies use highly artificial rhythmic stimuli, so it is unclear whether these findings generalize to real-world music and speech, which contain considerable tempo variability. Second, real-world interpersonal entrainment is typically bidirectional, but most neural entrainment studies examine an individual's neural entrainment to a fixed preprogrammed stimulus. We measured body sway and EEG in two professional string quartets ($n = 8$ musicians) as a real-world example of dynamic sensorimotor interpersonal interaction. We experimentally manipulated leadership, assigning a different musician as leader on each short performance. To investigate the direction and the magnitude of interpersonal entrainment among performers at both the behavioural and neural levels, we used Granger causality, in which the time series of movements or EEG from one musician are used to predict the upcoming movements or EEG from a second musician over and above prediction within a time series. We hypothesized that (1) leaders would influence others more than followers, and (2) the overall magnitude of interpersonal entrainment would predict the quality of the performance on a trial-by-trial basis. Body sway analyses revealed that, as predicted, leaders influenced follower more than vice versa, and the higher the total entrainment strength among the performers, the better the performance quality. For the EEG analyses we used a beamformer to reconstruct the source waveforms generated from auditory cortex, visual cortex, supplementary motor area (SMA) and frontal regions (DLPFC & VLPFC). Initial analyses showed that within-individuals, connectivity was stronger from auditory and visual cortex to SMA and frontal regions than in other directions, which is consistent with the view that sensory information guides executive control (frontal) and motor planning (SMA). Furthermore, across musicians, leader's neural oscillations across all regions better predicted follower's auditory neural oscillations than vice versa. Thus we have shown that interpersonal entrainment can be represented in neural oscillations and reflect directional influences between people.

68. Temporal Tracking of Speech Periodicity in Human Auditory Cortex

Ning Guo, Xiaopeng Si, Wenjing Zhou, Bo Hong

Periodicity is an important temporal feature of natural sound, and it is also the signal basis for the auditory perception such as pitch, semantic tone and intonation (Rosen, 1992). How periodicity of speech is parsed in the auditory system is an important topic of auditory neural coding. A large number of imaging studies have found that sounds with periodic structure elicits prominent activity in human primary auditory cortex (Zatorre, 2000; Patterson, 2002) . Electrophysiological studies of animal and human brain have found that the temporal pattern of the auditory neural responses also encodes periodic information (Wang, 2012; Nourski,2013). The periodic nature of speech, especially how the fundamental frequency f_0 is processed in the human brain, remains controversial (Wang, 2013). In this study, 30 epilepsy patients implanted with intracranial electrodes for clinical diagnosis passively listened to a set of periodic synthesized sounds and speech materials. By analyzing the intracranial EEG (iEEG) responses of 2722 electrodes, we identified xx electrodes over auditory cortex with responses tracking the fundamental frequency in the sounds. With a set of f_0 sweep stimuli, we observed that the tracking response of intracranial EEG is limited to the frequency band of 60-150 Hz, which partially coincides with the fundamental frequency distribution of the human speech, suggesting that temporal following is one of the possible mechanisms for encoding the fundamental frequency of speech in the human auditory cortex. These temporal following responses were mainly located in the Heschl's Gyrus(HG), planum temporale (PT) and the posterior part of the superior temporal gyrus (STG). The strength of temporal following (measured by phase locking value) of iEEG over the PT area on both hemispheres are significantly higher than that of the STG (Wilcoxon rank-sum test; $p < 0.05$). Compared with pure tone stimuli, PT and STG on both hemispheres were more responsive to harmonic stimuli and speech with similar fundamental frequency (t-test; $p < 0.01$). We also compared the temporal following strength of left and right auditory cortex, but no significant lateralization was observed.

69. Descending projections from the auditory cortex and inferior colliculus contact GABAergic cells in the ventral nucleus of the trapezoid body

Nichole L. Beebe, William A. Noftz, Brett R. Schofield

The ventral nucleus of the trapezoid body (VNTB), an auditory nucleus in the superior olivary complex (SOC), is a major target of descending projections from the auditory cortex (AC) and inferior colliculus (IC). Previous investigators reported that these descending projections contact cholinergic olivocochlear cells in the VNTB (Mulders and Robertson, 2000; Suthakar and Ryugo, 2017). The VNTB also contains GABAergic cells, which can project to the cochlear nuclei or inferior colliculus. Here, we asked whether GABAergic cells of the VNTB are contacted by descending projections from the AC and the IC.

We injected traditional tracers (e.g. FluoroRuby) or adeno-associated viruses carrying fluorescent protein genes (e.g., AAV2-hSyn1-EYFP) into the AC and/or IC of pigmented guinea pigs and Long-Evans rats. After 5-28 days, we perfused the animals and stained brain sections for glutamic acid decarboxylase (GAD) to label GABAergic cells.

Injections of anterograde tracer into the IC labeled axons and boutons in the thalamus, nuclei of the lateral lemniscus, SOC, and cochlear nucleus. In the VNTB, we saw many putative contacts onto GAD+ cells in both guinea pigs and rats. Injections of anterograde tracer into the AC labeled axons and terminals in the auditory thalamus, IC, and SOC. In the VNTB, putative contacts were present on GABAergic VNTB cells in both guinea pigs and rats, although these contacts were fewer than those observed after IC tracer injections. In animals with injections of different tracers into AC and IC, we observed convergence of AC and IC inputs onto single GABAergic cells of the ipsilateral VNTB.

These results show that descending projections from the AC and the IC contact VNTB GABAergic cells. Both AC and IC projections arise from excitatory cells, so their targeting of VNTB GABAergic cells could provide for top-down inhibition of nuclei innervated by the VNTB. Moreover, the convergence of AC and IC projections onto individual VNTB cells suggests integration of these descending inputs. Thus, the VNTB is well-situated to act as an inhibitory hub of the descending auditory system. It is possible that activation of these descending pathways contributes to habituation or attentional mechanisms to suppress or facilitate processing of auditory stimuli based on salience.

Supported by NIH R01 DC004391.

70. Distinct contributions of auditory cortical inhibitory neuron types to spectral surround suppression

Anna Lakunina, Yashar Ahmadian, Santiago Jaramillo

Spectral surround suppression in the auditory cortex is the phenomenon in which a neuron's response decreases when sound power increases in frequencies outside the neuron's classical receptive field. This process is potentially useful for filtering out broadband background acoustic noise. The role of different sources of cortical inhibition in generating spectral surround suppression remains unclear. Here, we evaluated the contributions to spectral surround suppression of the two most common inhibitory cell types in the cortex, parvalbumin-expressing (PV) and somatostatin-expressing (SOM) interneurons. First, we used optogenetic tools to tag PV and SOM cells in the auditory cortex and identify them during extracellular recordings in awake adult mice. We characterized the responses of PV, SOM, and putative pyramidal neurons to amplitude modulated bandpass filtered noise spectrally centered on the cells' preferred frequency. By varying the bandwidth of the stimulus, we determined the effect of adding power to the spectral surround on the responses of these cells. We found that both PV and SOM cells exhibit little spectral surround suppression compared to pyramidal neurons. In addition, most SOM cells respond over the entire duration of the sound, and have a substantially greater sustained response compared to PV cells. Next, we tested the causal role of PV and SOM cells in mediating spectral surround suppression by measuring the effect of inactivating each cell class on the responses of auditory cortical neurons to bandpass filtered noise. We found that inactivation of SOM cells, but not PV cells, significantly reduced the spectral surround suppression observed in auditory cortical neurons. These results suggest that SOM cells provide sustained, sound-driven inhibition that mediates spectral surround suppression in the auditory cortex.

71. Classifying the neural code of concurrently presented vowels

Samuel Smith, Ananthakrishna Chintanpalli, Mark Wallace, Adam Hockley, Michael Heinz, Christian Sumner

Listeners use multiple cues to help them understand speech despite interfering sounds (e.g. pitch, onset/offset asynchronies, dynamics). Historically, special attention has been paid to fundamental frequency (F0) cues and how they benefit the task of identifying concurrently presented vowels, denoted the 'F0 benefit' (review: Micheyl and Oxenham, 2010; *Hear. Res.* 266: 36-51). The classical model of concurrent-vowel identification can replicate an 'F0 benefit' only qualitatively (Meddis & Hewitt, 1992; *JASA* 91: 233-245). The model's utilisation of F0 cues requires temporal processing that is unverified by neurophysiological studies. Moreover, the model makes deterministic decisions resulting in very poor predictions of listener confusions. Significant flaws in this model, and within others, imply we do not understand how listeners solve this simple instance of the 'cocktail party problem'.

We present our model of concurrent-vowel identification, based on ideal observer principles. Our model takes expected neural responses to concurrent-vowel pairs, generated from a 'linear' simulation of the auditory nerve, as input. Likelihood functions predicting neural responses to each vowel combination are stored as templates. A high dimensional naïve Bayes classifier optimally compares raw input to these stored templates, and produces a confusion matrix. Our model can qualitatively and quantitatively replicate an 'F0 benefit'. This is despite having no explicit temporal processing of F0 cues. Simply, when there are larger F0 differences between vowels there is more information available for classification. Additionally, the model makes quantifiable predictions of listener confusions with a high degree of accuracy ($R > 0.9$). The model's generality allows it to also handle data produced by a 'nonlinear' simulation of the auditory nerve, or neural responses recorded from multi-unit electrodes in the guinea pig inferior colliculus. Overall, our model is much closer to predicting human performance at the concurrent-vowel identification task than previous models. The work promotes the use of ideal observer principles to develop our understanding of how listeners solve the 'cocktail party problem'.

72. A potential pathway mediating sound-induced anxiety-like behaviour

Yiwei Wang, Miaomiao Liu, Jingshan Zhou, Dongqin Cai, Limei Ke, Ling You, Fenghua Xie, Yin Yue, Kexin Yuan

Anxiety behavior is essential for animal to fight against potential threats in their environment. However, the neural circuits responsible for sound-induced anxiety-like behavior remain largely unclear. As an important division of medial hypothalamic defensive system, the ventromedial hypothalamus (VMH) is closely related to animal innate defensive behavior. Since it was reported that VMH receives direct inputs from non-lemniscal auditory thalamus, we propose that the non-lemniscal auditory thalamus→VMH pathway might play an important role in sound-induced anxiety-like behavior. Optogenetic activation of the non-lemniscal auditory thalamus→VMH fibers was sufficient to initiate anxiety-like behavior, such as a dramatic increase in thigmotaxis, rearing and mice respiration rate. Open field test showed that mice spent less time in the center of the chamber, and their average velocity as well as center entries decreased after laser stimulation. Elevated plus maze test exhibited that mice spent significantly longer time in closed arm during laser stimulation. In anesthetic mice, selective photo-activation of non-lemniscus auditory thalamus→VMH pathway increased respiration rate. Moreover, fiberphotometry recording showed considerably stronger responses of VMH neurons receiving direct non-lemniscal auditory inputs to threatening sounds compared with responses to white noise. This is the first study demonstrating sound-driven responses in mammalian VMH, and suggested that, besides non-lemniscal auditory thalamus→amygdala pathway, the non-lemniscal auditory thalamus→VMH pathway might play a role in sound-induced defensive behavior as well. Further loss-of-function studies would be necessary to consolidate this conclusion.

73. Cortical excitatory and inhibitory neurons differentially affect collicular responses to sound

Jennifer M. Blackwell, Winnie Rao, Dashiell Ridolfi-Starr, Maria N. Geffen

Within the auditory cortex, excitatory and inhibitory dynamics determine frequency selectivity in excitatory cells (Hamilton et al., 2013, Aizenberg et al., 2015; Seybold et al., 2015; Natan et al., 2015; Phillips et al., 2016). Auditory cortex sends an extensive descending pathway to the inferior colliculus, but how excitatory-inhibitory dynamics in the cortex affect this pathway remains unknown. Previous studies demonstrated that responses of neurons in the inferior colliculus are altered by focal electrical stimulation and pharmacological inactivation of auditory cortex (Jen et al., 1998; Yan et al., 2005; Zhang et al., 2005), but these methods lack the ability to manipulate specific cell-types. In this study, we use optogenetic techniques to modulate activity of excitatory cells, parvalbumin-positive (PV) interneurons, and somatostatin-positive (SOM) interneurons in the auditory cortex to test whether and how this pathway modulates frequency selectivity in the inferior colliculus in both anesthetized and awake mice. As expected from previous studies, activation of cortical excitatory cells increased spontaneous activity and decreased frequency selectivity, while activation of PV interneurons and SOM interneurons decreases spontaneous activity and increased frequency selectivity in putative excitatory cells in the auditory cortex. Furthermore, suppression of PV interneurons had the opposite effects on putative excitatory cell activity. However, whereas activation of excitatory cells decreased frequency selectivity and tone-evoked responses in the inferior colliculus, modulation of PV and SOM interneurons had weak effects on activity in the inferior colliculus. These findings suggest that modulation of frequency selectivity in auditory cortex by inhibitory neurons does not necessarily propagate to the inferior colliculus and that PV and SOM interneurons may differentially affect activity in the inferior colliculus.

74. Front-Temporal Cortical Interactions during Vocal Production in Marmoset Monkeys

Joji Tsunada, Steven Eliades

Vocal production, including human speech, is a sensory-motor process that requires self-monitoring of produced vocalizations to correct errors between intended and actual outputs. Recent evidence has demonstrated a suppression of auditory cortex during vocal production, a sensory-motor process that is theorized to involve efference copy signals originating in frontal cortical areas. However, whether or not there is communication between frontal or motor areas and auditory cortex during vocalization is unknown. Here, we simultaneously recorded neural activity from both auditory and frontal cortices of marmoset monkeys while marmosets produced self-initiated vocalizations. We found modulations of neural activity in both brain areas immediately before and during vocal production. Interestingly, theta-band activity, thought to be involved in coordinating neural activities, was observed to increase in both brain areas immediately prior to the onset of vocal production. We further tested the timing relationship between activity in auditory and frontal cortex. We found a subset pairs of recording sites with temporally covarying activities, including frontal activation preceding auditory just prior to vocal production, followed by auditory activation that preceded frontal once vocalization began. These results suggest communication between cortical areas during vocal production, with frontal-auditory pre-vocal signals that may reflect preparatory activity, and auditory-frontal signals that may represent self-monitoring of vocal feedback. These different neural interactions between auditory and frontal cortices may underlie mechanisms to calculate and correct for errors between intended and actual vocal outputs during vocal production.

75. Characterization of auditory neurons in the mouse superior colliculus detected using virtual auditory space stimulation

Shinya Ito, David A. Feldheim, Alan M. Litke

The superior colliculus (SC) is a midbrain area where sensory information is integrated to initiate motor commands. A number of advances in mouse molecular genetics, large-scale physiological recordings and SC-dependent visual behavioral assays have made the mouse an ideal model to understand the relationships between genetics, circuits and behavior in the visual system. However, the understanding of how auditory information is processed in the mouse SC is limited.

An important property of the auditory neurons in the SC is their tuning to sound source location. The topographic auditory spatial map in the deep SC has been characterized in species such as the ferret and barn owl, but not in mice. Therefore, we developed a system that presents auditory stimulation in a virtual space and recorded the response properties of deep SC neurons. We first measured the head-related transfer function—the sound modulation produced by the head and ears before the sound vibrates the ear drums—as a function of incident sound source directions and used this to produce the sound stimulus consistent with a specific source direction. This system allows flexible control of the stimuli and is compatible with large-scale recording systems that often create acoustic obstacles.

Using this experimental system we find neurons in the deep SC that are tuned to a specific (virtual) sound source direction. The properties of the auditory receptive fields and the topographic map will be presented. These results demonstrate that the mouse is a model to study mechanisms of auditory circuit formation and function.

76. O-15 Water PET study of speech in noise processing in cochlear implant patients

Laurie Ponto, Inyong Choi, Bruce Gantz, Bob McMurray, Tim Griffiths

One of the most important issues in hearing impairment (HI) is difficulty with speech in noisy real-world environments. Recent research in normal hearing listeners indicates that auditory cortex is active while abstracting speech objects from noise and provides input to fronto-temporal networks for further perceptual, attentional, and semantic analysis. We sought to understand whether the neural mechanisms by which cochlear implant (CI) listeners detect speech in noisy environments are similar to individuals with normal hearing. We tested a group of hearing-preservation (hybrid) CI users with devices that combine residual low frequency acoustic hearing with high frequency electric hearing.

N = 9 CI users were compared to N = 10 age-matched normal-hearing (NH) control participants. Cerebral blood flow was measured using [^{15}O]Water positron emission tomography while the participants listened to 2-min blocks of continuous sentences-in-noise or noise alone (matched on RMS sound level). On a given run for speech-in-noise (+7 dB), 30 unique sentence tokens (~2.5 sec length) were presented (1.5 sec inter stimulus interval). Twelve scans (6 each condition, random order) were acquired to allow for single-subject inference.

Robust activations were found in single subjects for the contrast speech-in-noise vs noise in the auditory cortex (lateral Heschl's gyrus, planum polare, and planum temporale) and inferior frontal cortex (frontal operculum and inferior frontal gyrus) bilaterally ($p < 0.05$). The activity patterns were very similar across subjects.

The NH control participants were compared to the CI users for evaluation of the neural substrate of the network and overall activity level. No significant interaction was demonstrated between group and the speech-in-noise minus noise contrast in auditory and inferior frontal cortex. Six NH control participants were also scanned for a second session within six months of the previous scan to evaluate the reliability of activation patterns. Across 78 anatomically-defined regions-of-interest, regional activation levels normalized to global activity were compared between time 1 and time 2 for both noise and speech-in-noise conditions. All correlations between the sessions were above 0.9.

A frontal and temporal network of brain activity was demonstrated for speech-in-noise in hybrid CI users. These results were reliable at the single-subject level, which allows for a longitudinal investigation of CI users after device activation. The results support a similar system for speech-in-noise processing in hybrid CI patients and normal hearing listeners.

77. Neural computational principles of auditory processing revealed by magnetoencephalography and deep neural networks

Xiangbin. Teng, David Poeppel

Acoustic information is processed by the auditory system in a hierarchical manner – detailed acoustic information is extracted and represented at the low level (e.g. sub-cortical structures and primary auditory cortex) while relatively abstract information of category and identity of sounds is formed at higher levels (e.g. superior temporal gyrus, superior temporal sulcus and frontal areas). Uncovering this hierarchical principle of information processing, neurophysiologically and computationally is of foundational significance. To approach this, we ask, neurophysiologically, how pitch and timbre information is extracted along the auditory hierarchy temporally and spatially using magnetoencephalography (MEG). We simulate, computationally, how deep neural network (DNN) represents such information across different layers. We selected three different musical instruments and 8 different notes from the Nsynth Database and presented each sound 100 times to listeners while recording MEG signals. We classified different musical notes and reconstructed each sound along auditory ventral pathway from reconstructed MEG source signals. Taking advantage of the large number of sound samples available to train a DNN from the Nsynth Database, we trained a DNN to classify different notes and, corresponding to the MEG paradigm, studied the representation of acoustic details and classification performance on each layer of the DNN. We hypothesize that the stimulus reconstruction performance decreases and classification performance increases along the auditory hierarchy; in parallel, in middle layers of the DNN, representation of detailed acoustic information becomes degraded but performance on classification of notes is improved on top layers of the DNN. By combining results from neurophysiology and DNN, we can begin to decipher auditory processes in the brain, temporally, spatially, and computationally.

78. Investigating auditory conscious perception with a threshold task and intracranial EEG

Kate L. Christison-Lagay, Christopher Micek, Sharif I. Kronemer, Sarit Forman, Mark Aksen, Ahmad Abdel-Aty, Frederick van Duyne, Melanie Boly, Elsa Juan, Tom Bugnon, Erin Yeagle, Jose Herrero, Stephan Bickel, Ashesh Mehta, Lawrence J. Hirsch, Jason L. Gerrard, Dennis D. Spencer, Hal Blumenfeld

Recent work has made significant headway in understanding the temporal and spatial dynamics of the neural mechanisms of consciousness. Our lab previously developed a perceptual threshold task that used a face stimulus calibrated to a subject's 50% detection rate (Herman et al. 2018, Cerebral Cortex). Subjects were subsequently prompted to report whether they perceived the stimulus, and the stimulus' location (to validate perception). We have used this visual paradigm in conjunction with intracranial electroencephalography (icEEG) in patients with intractable epilepsy and found initial activation of early visual cortex for both perceived and not perceived stimuli. In perceived trials only, this was followed by a decrease in both the early visual areas and the default mode network and a wave of activity that swept through the cerebral cortex, followed by a late reactivation of the early visual areas. We call this pattern of activity the "switch-and-wave." In the current study, we expand our investigation of conscious perception to the auditory domain to establish whether similar "switch-and-wave" activity accompanies conscious auditory perception. We have developed an auditory counterpart to our visual task, in which three target sounds are calibrated to a subject's 50% detection threshold and embedded in white noise. Each target sound (a whistle, a 'laser', and a waterdrop) are presented for 75ms, and subjects are prompted to report whether they perceived the sound, and the sound's identity (to validate perception); when participants indicate they do not hear a sound, they are asked to randomly guess on the second forced-choice question about sound identity. In behavioral studies (n=25), we find that the stimulus perception rate was 57.5% (2.7% SEM) when the target was present whereas false positive rate was 10.9% (2.9% SEM) for blank trials. When subjects indicated they heard a target sound, they correctly indicated the sound's identity in 89.5% (2.3% SEM) of trials; when they indicated they had not heard a sound, sound identification accuracy was 36.7% (2.2% SEM) (chance is 33%). Therefore, this represents a robust behavioral paradigm for testing conscious auditory perception. We have begun testing this auditory paradigm in epilepsy patients with icEEG at multiple centers. Early analyses of broadband gamma power suggest there are similarities in the conscious perception of auditory and visual stimuli. Analyses of broadband gamma power in relation to this task are ongoing.

79. Removable cranial window for chronic multi-scale and multi-modal optical imaging in marmosets

Yueqi Guo, Xindong Song, Xiaoqin Wang

The common marmoset (*Callithrix jacchus*), a highly vocal New World monkey species, has garnered considerable interest in recent years as a promising primate model in neuroscience studies. The smooth cortical surface of the marmoset brain provides a special advantage for studying cortical functions using optical imaging methods.

In this study, we developed procedures to perform multi-scale and multi-modal optical imaging in the auditory cortex of awake marmosets. The major goals of our procedures are: 1) targeted neuron labeling with virus-mediated methods, based on data obtained with intrinsic imaging; 2) a removable imaging window design, allowing us to remove and replace the window for different imaging modalities.

To perform intrinsic imaging and virus injection, we designed an artificial dura (AD) based cranial window. Since the dura of marmosets is opaque and must be removed for optical imaging, an AD is necessary for protecting the cortex, preventing the dura growing back and thus keeping the window clear. The optically flat AD is made of silicone (Shin-Etsu, ET-1300T), with a diameter of a quarter inch, and thickness of 150~200um. An additional 1.3mm wall structure along with a 3mm wide flange that was inserted under the surrounding dura, prevent the dura growing back into the imaging window. A customized rubber chamber structure was used to cover and protect the window when the animal is back to its cage.

After obtaining a functional map with intrinsic imaging through the chronically-implanted window, we were able to achieve targeted GCaMP-carrying virus injection, which allowed us to perform calcium fluorescence imaging afterward. The virus was mixed with the dye (Fast Green) to visualize the diffusion pattern and possible back-flow when injecting the virus. The mixture was loaded into glass micropipettes, which were pulled and ground so that the tip has a ~30um outer diameter and a ~20° angle. We were able to penetrate the silicone AD easily and achieve consistent injection speeds by manually pushing a syringe air-coupled with the micropipette.

To achieve higher brain stability during two-photon imaging in awake marmosets, the original silicone based cranial window was removed and replaced by a modified AD, which has a coverslip in the middle. Through this modified window we were able to record calcium signals from individual neurons with a customized two-photon microscope.

The removable cranial window was safe and relatively easy to maintain. It enabled chronic optical imaging in marmosets over several months.

80. Hearing loss, auditory sensorimotor gating deficits, and cortical interneuron abnormalities in a mouse model of 22q11.2 Deletion Syndrome

Jennifer F. Linden, Fhatarah A. Zinnamon, Freya G. Harrison, Sandra S. Wenas, Arne F. Meyer, Qing Liu and Kuan H. Wang

The chromosomal microdeletion that causes 22q11.2 Deletion Syndrome (22q11DS) confers a 25-30% risk of developing schizophrenia. An estimated 40-70% of patients with 22q11DS also have hearing loss, mostly from otitis media (OM). In the Df1/+ mouse model of 22q11DS, similar rates of hearing loss and OM occur (Fuchs et al. PLoS One 2013) and have been shown to arise from haploinsufficiency of the TBX1 gene within the minimum 22q11DS deletion region (Fuchs et al. Hum Mol Genet 2015). However, the relationship between hearing loss and susceptibility to schizophrenia-relevant brain and behavioral abnormalities in 22q11DS is unknown. Here we report a link between hearing loss, deficits in auditory sensorimotor gating, and abnormalities in parvalbumin-positive auditory cortical interneurons in a mouse model of 22q11DS.

Using the auditory brainstem response (ABR), we found that 60% of Df1/+ mice had hearing loss in one or both ears. However, suprathreshold cortical auditory evoked potentials (AEPs) were similar in Df1/+ and WT mice. The ratio between AEP P1-N1 or N1-P2 amplitude and ABR wave I amplitude was significantly higher in Df1/+ mice with hearing loss than in WT mice or Df1/+ mice without hearing loss, suggesting a compensatory increase in central auditory gain in Df1/+ mice with hearing loss.

Behavioural measures similarly revealed an influence of hearing loss. Acoustic startle response (ASR) thresholds were significantly higher in Df1/+ than WT mice. Prepulse inhibition (PPI) of ASR was reduced in Df1/+ mice relative to WT for prepulse cues with fixed sound level, as has previously been reported elsewhere. However, we found that there was no significant difference in PPI between Df1/+ and WT mice when the prepulse cue sound level was adjusted to be constant relative to the startle threshold for each animal.

Finally, in immunohistochemical studies, we found that the density of parvalbumin-positive (PV+) interneurons in the auditory cortex was significantly reduced in Df1/+ compared to WT mice. This abnormality arose primarily in Df1/+ mice with hearing loss, suggesting cortical compensation for loss of input.

Overall, the findings indicate that Df1/+ mice have reduced hearing sensitivity and elevated startle thresholds, but also increased central auditory gain and reduced density of PV+ inhibitory interneurons in auditory cortex. Moreover, PPI of acoustic startle in Df1/+ and WT mice differs for prepulse cues of fixed sound level, but not when the cue level is adjusted relative to startle threshold. Thus, the findings suggest a complex interaction between hearing loss and auditory brain and behavioural abnormalities in 22q11DS models.

Supported by: Action on Hearing Loss, National Institute of Mental Health intramural research program

81. Receptive field characterization using pure tones and broadband stimuli in the mouse primary auditory cortex

Kyunghee X. Kim, Craig A. Atencio, Christoph E. Schreiner

The primary auditory cortex (A1) is a hub that communicates with subcortical and cortical regions. Sound information transferred from the medial geniculate body is further modified via intracortical horizontal pathways in A1. Although detailed anatomical pathway information is needed, functional descriptions of receptive field properties unambiguously illustrate spectral processing. Spectral integration in A1 can be shown via intracellular higher resolution recordings of the subthreshold receptive field. While subthreshold responses with intracellular recordings in A1 have been measured with pure tones, there is little evidence that compares the detailed properties of receptive fields with subthreshold and spiking responses using both pure tone and complex stimuli. We used anesthetized female C57BL/6 mice between 4 and 11 weeks of age and performed in vivo blind patch-clamp recordings in current clamp mode from A1 neurons located at a depth of 300–500 μm by applying pure tone (4–40 kHz, 0–70 dB) and broadband (0.5–40 kHz, 50dB relative to the pure tone threshold) dynamic moving ripple (DMR) stimuli. We constructed the tonal receptive field (TRF) from responses with pure tones and computed the spectrotemporal receptive field (STRF) using the spike-triggered average (STA) analysis for responses with the DMR stimulus. Sharpness of tuning ($Q = \text{characteristic frequency} / \text{bandwidth}$) for TRFs and STRFs was measured. TRFs from spiking responses showed narrower tuning than TRFs from subthreshold responses. STRFs in both subthreshold and spiking responses were more sharply tuned than TRFs. We conclude that single neurons in A1 perform spectral integration over a broad range of frequencies. The excitatory/inhibitory weighing for each neuron's spectral integration is modified by dynamic cortical activity resulting from inputs of spectrotemporally complex stimuli.

82. Opposite abnormalities in gap-in-noise sensitivity in the auditory midbrain and thalamus of a mouse model of developmental disorder

J. Mattley, L.A. Anderson, J. F. Linden

The BXSB/MpJ-Yaa mouse is a powerful animal model for studying the neural mechanisms of gap-detection deficits. Around 30-50% of BXSB/MpJ-Yaa mice have small cortical malformations called ectopias. Previous work has shown that ectopic mice have more difficulty detecting very short gaps in noise than non-ectopic mice (Clark et al. 2000). Furthermore, minimum gap durations required to elicit significant changes in the activity of auditory thalamic neurons are longer in ectopic than non-ectopic mice (Anderson and Linden 2016). Moreover, sound-offset responses --- transient increases in firing at sound termination --- are less common in auditory thalamus of ectopic mice, suggesting the hypothesis that the abnormal thalamic gap-in-noise sensitivity might arise from a deficit in sound-offset responses (Anderson and Linden 2016).

To determine whether neural deficits in gap-in-noise sensitivity and sound-offset responses in the auditory thalamus of ectopic mice are inherited from the midbrain, we made extracellular recordings from the inferior colliculus (IC) in 10 ectopic and 10 non-ectopic BXSB/MpJ-Yaa mice. Mice were anaesthetised with urethane (as in the previous thalamic study), and neural responses recorded with 16-channel microelectrode arrays (Neuronexus). In contrast to the previous results from the lemniscal (primary) auditory thalamus, in IC we found that minimum gap durations for evoking responses to gap-in-noise stimuli were shorter in ectopic than non-ectopic mice, for neurons with V-shaped tuning curves typical of the lemniscal IC (rank-sum $p < 0.001$; ectopic $n = 133$ recordings, non-ectopic $n = 79$). However, in agreement with previous thalamic results, we found that the proportion of cells with sound-offset responses in IC was significantly reduced in ectopic mice (9% in ectopic, 15% in non-ectopic across all IC recordings; Fisher's exact test $p = 0.03$). These results indicate that the offset-response deficit in auditory thalamus of ectopic BXSB/MpJ-Yaa mice may be inherited from the IC. Moreover, the observation that in ectopic mice, minimum gap duration thresholds are abnormally high in auditory thalamic neurons but abnormally low in IC neurons suggests a critical role for midbrain and/or thalamic circuitry in development of gap-detection deficits.

References:

- Anderson LA, Linden JF (2016). Mind the gap: two dissociable mechanisms of temporal processing in the auditory system. *J Neurosci* 36:1977-95.
- Clark MG, Sherman GF, Bimonte HA, Fitch RH (2000). Perceptual auditory gap detection deficits in male BXSB mice with cerebrocortical ectopias. *Neuroreport* 11:693-96.

83. Co-variability across the neural population in the map of auditory space of the barn owl

Michael V. Beckert, Roland Ferger, Keanu Shadron, Brian J. Fischer, Jose L. Pena

The midbrain of the barn owl contains a map of auditory space. However, the characterization of the map has been based mainly on single-cell or local multiunit recordings. This approach is blind to co-variabilities in response strength and spike timing that occur on a trial-by-trial bases. However, these correlations have profound effects on information encoded by neural populations, depending on their structure as well as the system they are embedded within. We therefore sought to expand upon previous work by measuring population activity in the map-of-space.

To this end we have developed multi-electrode array recordings, utilizing a microdrive capable of positioning electrodes deep within the brain, allowing access to the midbrain map. Using this approach, data were collected to quantify the similarity of tuning (signal correlation), trial-by-trial co-variability (noise correlation), as well as precision of spike timing (spike-time synchrony).

As anticipated, more distant neurons displayed weaker signal correlations. Additionally, noise correlations were directly related to the signal correlation between pairs of neurons, similar to observations in the visual system. On the other hand there was no relationship between the level of spike-time synchrony and signal correlation. We assessed the impact of these correlations on information by training a decoder with neural data, which preserved the natural correlations. We compared this to a decoder that removes correlations by shuffling the trials. This work expands upon previous work that investigated the sound localization system of the barn owl using either recordings from single neurons or multiple nearby neurons.

84. The primate voice-sensitive cortex is a general acoustic pattern analyser

Matthias Staib, Sascha Frühholz

The perception of human voices activates extended regions of the auditory cortex known as temporal voice areas (TVA). However, it is unclear whether the observed activity in the TVA represents encoding of the acoustic features that discriminate voices from non-voices, or whether these activation patterns are shared across acoustic domains, such as recorded voices or synthetic sounds. Using a reconstruction model for fMRI data, we demonstrate in 25 healthy human volunteers that the spatial configuration of the activation pattern in primary and higher auditory cortices that discriminates listening to voices from non-voices (target pattern) can be reconstructed with high accuracy from activation patterns elicited by listening to synthetic sounds (reconstruction patterns). The reconstruction model assigns a linear weight to each reconstruction pattern such that the target pattern is fully explained. We found that the linear weights are associated with the acoustic features of the corresponding synthetic sounds across auditory areas. Importantly, in higher auditory areas, the weights can be predicted from each participant's rating of voice-similarity for the synthetic sounds over and above their acoustic features. Using a decoding approach with cross-classification, we then confirm that perceived voice-similarity of recorded and synthetic sounds generalizes across acoustic domains. Our results show that specific activation patterns in the auditory cortex that encode vocal and non-vocal categories represent voice-similarity tracking for sounds across a vast acoustic space. Overall our findings suggest, first, that several auditory areas are involved in estimating voice-similarity across a wide range of acoustic stimuli, and second, that voice-similarity is represented by an invariable activation pattern in higher auditory brain areas.

85. Cortical Asymmetry for Sound Energy (CASE) in Mustached Bat Primary Auditory Cortex

Stuart D. Washington, Georgios A. Keliris, Jagmeet S. Kanwal

Mustached bats emit a wide variety of conspecific social calls composed of tonal and frequency modulated (FM) signals. Their left hemispheric dominance for processing social calls in primary auditory cortex (A1) is linked to computational strategies mirroring the asymmetric sampling of time (AST) hypothesized to underlie left hemispheric dominance for speech in humans. However, this bat-based AST is more evident in males than females. Here, we report Cortical Asymmetry for Sound Energy (CASE), a computational strategy that coexists with AST in bats. We systematically varied the amplitudes (1-91 dB SPL) of sound stimuli while recording A1 neural responses from the left (n=179 neurons) and right (n=147 neurons) hemispheres of 10 (four female and six male) bats. We used two sets of stimuli: a constant, 30 ms tone and a frequency modulated (FM) signal. The tone was optimized to the best frequency of each neuron (BF, ~57-60 kHz) and the FM signal was optimized for rate, bandwidth, direction, center frequency, and duration based on each neuron's selectivity. Right A1 neural responses had longer latencies than left A1 responses when elicited by best-FMs (left: 14.82 ms \pm 0.61 s.e.m.; right: 27.35 ms \pm 2.34 s.e.m.; $p = 8.08 \times 10^{-7}$). These were drastically reduced for the tonal signals (left: 14.92 ms \pm 0.35 s.e.m.; right: 16.08 ms \pm 0.43 s.e.m.; $p = 0.03$). Right A1 neurons are generally selective for FMs with longer durations than left A1 neurons (left: 24.70 ms \pm 3.0 s.e.m.; right: 50.29 ms \pm 5.45 s.e.m.; $p = 6.13 \times 10^{-5}$), and the longer latencies amongst right A1 neural responses to FMs likely reflect the distribution of acoustic energy during the FM signal. The basis of this assertion is that amplitudes eliciting peak responses from A1 neurons for fixed-duration tones at best frequency were greater for the right hemisphere than for the left (left: 56.9 dB SPL \pm 1.06 s.e.m.; right: 64.6 dB SPL \pm 0.85 s.e.m.; $p = 3.14 \times 10^{-8}$), but this relationship fails for similar amplitudes of best-FMs (left: 53.07 dB SPL \pm 1.98 s.e.m.; right: 54.10 dB SPL \pm 1.93 s.e.m.; $p = 0.71$). Since neither latencies ($F [1, 244] = 0.43$; $p = 0.51$) nor tonal best amplitudes ($F [1, 298] = 0.25$; $p = 0.62$) were significantly affected by an interaction between hemisphere and sex, we conclude that CASE is independent of sex. The CASE computational strategy may underlie a previously reported right A1 selectivity for high intensity social calls in bats, a right hemispheric advantage for intensity discrimination in humans, and right hemispheric selectivity for long-duration sounds in multiple species.

86. Neuromodulation and Plasticity for a Rodent Model of Cochlear Implant Use

Erin Glennon, Jasmin Multani, Ioana Carcea, Mario Svirsky, Robert C. Froemke

Cochlear implants are neuroprosthetic devices that can provide hearing to deaf patients. However, learning rates and peak performances of speech perception with cochlear implants are highly variable (Blamey et al. 2013). Adaptation to cochlear implants is believed to require neuroplasticity within the central auditory system (Fallon et al. 2009). However, mechanisms by which behavioral training enables plasticity and improves outcomes are poorly understood. Here we investigate the hypothesis that neural mechanisms that promote plasticity in the rodent auditory system are key to optimizing cochlear implant usage, and might be especially helpful in cases of poor performance. We focus on noradrenergic modulation of rat auditory cortex by the locus coeruleus, which can enable robust and long-lasting neural and behavioral changes (Manunta and Edeline 2004; Martins and Froemke 2015; Sara 2015)

We developed a new surgical approach for cochlear implantation in adult rats (King et al. 2016). Our approach optimizes insertion depth of an 8-channel electrode array and allows rats to freely behave while using the implant to perform auditory tasks. Normal hearing rats are trained on a go/no-go task, and self-initiate trials to respond to a target tone. Previously, we showed that this task requires auditory cortex, and that this task is sensitive to cortical modulation and plasticity (Carcea et al. 2017).

We tested if changes in locus coeruleus activity affect or improve auditory learning in normal hearing and cochlear-implanted rats. Prior to each daily behavioral training session, rats underwent a 5-10 min locus coeruleus pairing session. We examined how locus coeruleus stimulation might affect reversal learning when a new sound became the target. The new target was paired with locus coeruleus stimulation as for cochlear-implanted animals. Locus coeruleus stimulation accelerated learning in each case. We used fiber photometry to monitor neural activity of noradrenergic locus coeruleus neurons, showing strong responses to novel auditory stimuli and noxious stimuli. These studies indicate that neuromodulation can play a powerful role in shaping outcomes with cochlear implant use and training.

87. Frequency dependent interaction of dual sound representations in monkey inferior colliculus

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

Behaving in complex environments requires the segregation and preservation of multiple, often overlapping stimuli. Neural codes relying on rate are potentially ill suited for the coding of multiple stimuli. A neuron cannot fire at two rates simultaneously. How, then, do neurons encode multiple stimuli that fall within their receptive fields? A useful structure to investigate this question is the primate inferior colliculus (IC); which possesses neurons with broad response functions in the spatial and frequency domains. Recent work showed some neurons in the monkey IC fluctuate between activity patterns that could permit encoding of simultaneously presented sounds across time (Caruso et. al., *BioRxiv*, 2017; Willett et. al., *Soc Neuro Abstr* 2016, Willett et. Al., *Soc Neuro Abstr* 2017). However, it remains uncertain what determines how a given neuron responds to a given combination of stimuli. The current study evaluates how the frequency separation between the pairs of sounds impacts the observed response pattern.

Single unit IC activity was recorded while rhesus macaques performed a dual sound localization task involving either a single saccade to the location of one sound or a sequence of two saccades to the location of each of two simultaneously presented sounds (bandpass noise of different center frequencies). We found that fluctuating activity patterns were more common on dual sound trials when the two stimuli were closer in frequency, and that when the stimuli were well separated, the responses more closely resembled the response to one of the two sounds when it was presented in isolation.

Together, these findings suggest that fluctuating activity contributes to coding when the stimuli individually evoke activity in a highly overlapping population of neurons, but when the stimuli are well separated, this is not necessary.

88. Voice activity detection in the brain: Understanding the spatio-temporal characteristics of voice detection in the auditory cortex during fMRI.

Huw Swanborough, Matthias Staib, Sascha Fruehholz

The ability to perceive speech-in-noise is vital to our ability to effectively communicate. However, despite the substantial body of work regarding speech-in-noise comprehension, there is a lack of understanding of the neurophysiological underpinnings of how the auditory system detects a voice within an auditory scene. This study works to address this gap by investigating the spatio-temporal characteristics of voice-activity detection in the brain, and how the affective quality of a voice may moderate this activity. To this end, the study approaches the auditory cortex(AC) under the assumption that discrete areas are analogous to a computerised voice activity detector, whereby voice-relevant, acoustic features are gradually accumulated from an auditory scene until a binary decision is made. During fMRI acquisition, participants were presented with ten-second samples of speech-in-noise at a continually increasing signal-to-noise ratio, participants responded when they were certain that they had detected voice activity present within the stimulus.

Using finite impulse response models, we compared the HRF characteristics in discrete areas of the AC and Amygdala, and how these were moderated by affect. We found evidence that specific populations of neurons in each participant, centred around Heschl's Gyrus, demonstrate a robust response to voice activity detection, but do not show a sustained BOLD response for the duration of voice listening. Additionally, we observed that different affective qualities of vocal stimuli result in differences in the temporal characteristics of the HRF for voice detection, as well observing possible evidence that could be interpreted as accrual of information prior to a perceptual decision being made. The findings from this study possibly indicate that speech-in-noise detection relies on 'voice relevant' acoustics properties reaching a statistical detection threshold, which may then initiate further voice-perception processes.

89. Neural substrates of auditory rhythm processing and language skill in early-to-mid adolescence.

Manon Grube, Faye Smith, Sukbinder Kumar, Heather Slater, Timothy D. Griffiths

This study seeks the neural substrates of auditory-sequence processing skills in the adolescent brain by seeking correlation between grey matter density and a systematic battery of tests to measure timing and rhythm skill. The study is part of a large initiative at a local high school where we carry out extensive behavioral testing to assess auditory skills and literacy skills on whole-year group cohorts (see PMID 22951739 for initial report). This study was carried out on subgroups from the larger study.

Previous work in 42 twelve-to-fourteen year-olds suggested a correlation between grey matter density in the left intra-parietal sulcus (IPS) and the first principal components of both auditory and language skills (Grube et al., Society for Neuroscience 2011: XX27 509.10). We also found a correlation between rhythm processing and grey matter density in the right cerebellum.

The current work assesses rhythm and language skill and their corresponding structural correlates of in two separate, new cohorts of mean age 12 (n = 20) and 14 (n = 24), respectively. Structural MRI was carried out at 3T on a Phillips xxx scanner, and voxel-based morphometry (VBM) implemented in SPM 8 sought correlation between grey matter density and auditory and language skill, taking non-verbal intelligence into account.

We test the hypotheses: i) there is a critical correlation between left IPS grey matter density and both sound sequence and literacy skill; ii) there is a correlation between cerebellar grey matter density and skill in rhythmic perception; iii) the alteration in the behavioral link between auditory sequencing and language skill from early to mid-adolescence that we have observed (Grube et al., Society for Neuroscience 2016: HHH3 85.07) is reflected in altered correlations between grey matter density and the sequence and rhythm measures.

90. Adaptation to simple sounds in the auditory field of insular cortex in anesthetized rats

The insular cortex is multi-modal cortical structure of mammalian brain composed of functionally distinct subregions characterized by different patterns of connectivity with other brain areas. Rat Insular cortex possess an auditory responsive field which is anatomically separated from other auditory cortical fields. While simple sound response properties of neurons in the insular auditory field have been described, we are not aware of any study of higher-order processing in this area. Stimulus specific adaptation (SSA) is the specific decrease in the response to a frequent ('standard') stimulus, which does not generalize, or generalizes only partially, to another, rare stimulus ('deviant'). Here we studied SSA in the insular auditory field. We used two approaches for recording neuronal activity in the insular cortex of halothane anesthetized rats. First, we characterized the location of insular auditory field in Sabra-R rats using widely spaced arrays of 9 single tungsten electrodes inserted vertically to various layers of the insular cortex and the claustrum, delineating the antero-posterior and medio-lateral coordinates of the field. Subsequently, we performed series of recordings using high-density silicone probes (neuropixels) to delineate dorso-ventral coordinates of the field and its microstructure. Both approaches revealed presence of the auditory responsive neurons in the insular cortex. Histological verification revealed that auditory responsive neurons were most frequently detected in the granular insular cortex between +0.24 mm to -1.08 mm posterior to bregma. Insular cortex neurons showed pronounced adaptation to repeated simple sounds and weak/moderate SSA. Consistent with the strong adaptation, these neurons failed the accepted test for deviance detection, showing responses to deviants that were substantially smaller than the responses to the same tones within sequences that contained multiple frequencies. Under anesthesia, the neuronal responses in the rat insular auditory field are therefore less capable of signaling deviance than neurons in primary auditory cortex.

91. Directed effective connectivity in the human and monkey brain: Auditory cortex impact on inferior frontal gyrus, hippocampus and anterior cingulate cortex

Francesca Rocchi*, Hiroyuki Oya*, Fabien Balezeau, Zsuzsanna Kocsis, Jeremy Greenlee, Timothy D. Griffiths, Matthew Howard III & Christopher I. Petkov

The correspondences or differences between human and nonhuman primates in connectivity between auditory cortex and other brain regions remain open questions: What is the evidence for evolutionary conservation or specialization in effective connectivity between auditory cortex and brain regions that in humans are implicated in language, memory or learned vocal production? Here we had the opportunity to directly compare electrical microstimulation of site-specific contacts available in human epilepsy patients being monitored for surgical resection. We also conducted comparative auditory cortex microstimulation in two rhesus macaques. Microstimulation was combined with functional Magnetic Resonance Imaging (fMRI) to assess the directed effective connectivity induced by electrically stimulating auditory cortical sites. Stimulation of primary auditory cortex (field A1) in the macaques resulted in significant fMRI responses in several brain regions ($p < 0.05$, cluster corrected), including the inferior frontal cortex (areas 44 and 45), regions of the hippocampus and the anterior cingulate cortex. In the monkeys, we also compared the stimulation of primary field A1 to sites in the caudal belt adjoining the non-primary fields CL and CM. In human patients with clinically implanted depth electrodes in Heschl's gyrus (HG), we microstimulated contacts located in the medial HG (a primary like region of auditory cortex; $n = 17$; 2 excluded from a total of 19 because of poor MRI registration). Stimulation of the medial HG resulted in significant fMRI responses in inferior frontal cortex, parts of the hippocampus and anterior cingulate cortex. These observations suggest a considerable amount of correspondence between humans and monkeys in effective connectivity between auditory cortex and other brain regions, which in humans have been implicated in various language functions, cognition or learned vocal production.

92. Mating but not restraint vocalizations have different valence for male and female mice

Zahra Ghasemahmad, Rishitha Panditi, Krish Nair, Bhavya Sharma, Jeff Wenstrup

Introduction:

Although studies have described the significance of mouse vocalizations from the sender's perspective, fewer have investigated effects of these vocal signals on the receiver. Here, we report playback experiments using natural vocal sequences from two behavioral contexts: mating and restraint. We examined whether these sequences altered the behavior of listening mice and whether male and female mice responded in the same manner.

Method:

Vocalizations and behavior of CBA/Cal mice ($n=12$; ages p90-p180) were recorded during the contexts of mating and restraint. We selected five exemplars of natural vocal sequences from restraint and from a higher arousal state of mating (behaviorally defined). Mating sequences included both male and female's vocalizations. These sequences were conditioned for use in playback experiments. Both male and female were used in behavioral tests. Only females in estrous were included in final data analysis. Prior to testing, mice experienced the behaviors of mating and restraint in a counterbalanced order on two consecutive days. On a subsequent test day, animals habituated to the test chamber for 3 hours, then were presented with either the mating or restraint vocalizations for 20 minutes. We video-recorded and analyzed 19 different behaviors before, during, and after stimulus presentation.

Results:

Mating and restraint vocal sequences both increased attending and head orientation toward the speaker ($p<0.0005$). In response to playback of vocal sequences linked to a high arousal stage of mating, females decreased locomotion, adopted an alert posture and escape behavior from the speaker; Males, however, increased exploratory behaviors (locomotion and rearing) in response to mating sequences and explored the speaker side more than females ($p<0.05$). In response to restraint vocal sequences, however, both males and females frequently approached the speaker area but not the speaker itself ($p<0.05$). Exposure to these sequences in both males and females resulted in escape behavior from the speaker (for both sexes, $p<0.0005$).

Conclusion:

Vocal sequences linked to mating change the behavior of listening male and female mice but do so in different ways consistent with their behaviors during mating: males are more exploratory, while females display reduced locomotion and more escape or alert behavior. Vocalizations linked to restraint context, however, result in an increase in escape behavior from the speaker for both sexes, as well as twitching and lung and sniff that are linked to the negative affect of

these vocalizations. Overall, our findings show that the valence of mating but not restraint vocal sequences differs based on the sex of the listening animals.

93. Nonlinear coding of naturalistic sound streams in marmoset auditory cortex

Luke A Shaheen, Stephen V David

The human auditory system is adept at isolating and comprehending a single sound source out of multiple sources (auditory streaming). The early auditory system encodes the incoming signal across a population of neurons, each sensitive to distinct low-level features, such as spectral frequency, spatial location, and modulation frequency. Because natural sound sources can contain overlapping features, the activity of individual neurons reflects a mixture of multiple sources. As signals pass through the auditory hierarchy, they undergo a series of nonlinear transformations which have been proposed to support emergent stream segregation. Such nonlinear responses are well-documented for static, synthetic stimuli in auditory cortex (ACtx). How completely the mechanisms identified in this work extend to natural, dynamic stimuli such as human speech is unknown.

To study streaming of naturalistic sounds at the single-neuron level, we recorded ACtx activity in passively-listening marmosets during presentation of a two-“voice” stimulus that retained the temporal dynamics of speech but had static spectra. Each voice consisted of a harmonic complex tone (HCT) with a unique fundamental frequency, modulated by a temporal envelope drawn from human speech. The envelope could be the same for both voices (coherent) or different (incoherent). Coherent voices were perceived as originating from a single source, and incoherent voices as two sources.

Responses to two dynamic HCTs were poorly predicted by responses to static HCTs, and by the sum of responses to each voice in isolation. Instead, responses were mostly suppressed relative to the sum of the isolated responses. However, periods of nonlinear enhancement were also observed. Nonlinear interactions varied with coherence. Greater suppression was observed for incoherent stimuli, possibly reflecting differences in streaming between conditions. A linear-nonlinear (LN) encoding model incorporating a spectro-temporal filter followed by a static nonlinearity was able to predict some nonlinear interactions (ex. saturation), but most neurons remained poorly predicted. Ongoing efforts will determine if a model incorporating nonlinear adaptation and/or multiple LN units is sufficient to predict responses in these conditions.

94. Auditory language localizer: An fMRI/ECoG study with epilepsy patients

B. F. Snoad, P. E. Gander, C. K. Kovach, K. V. Nourski, H. Kawasaki, E. Fedorenko, M. A. Howard III

Language processing is supported by an extensive network including temporal, frontal and parietal regions. Recent electrocorticographic (ECoG) investigations during language-based tasks have provided new insights in functional localization of mechanisms that support language processing. We set out to expand this work using a novel passive auditory language localizer task (Scott et al., *Cogn. Neurosci.* 2017, 8:167-76) to identify key language cortical regions in surgical epilepsy patients. The 16 pairs of clear and degraded speech monologues presented across two runs. We collected non-simultaneous functional magnetic resonance imaging (fMRI) and ECoG data in each subject.

The speech vs. silence contrast in fMRI data showed distinct activations in bilateral auditory cortex and planum temporale ($p < 0.05$, uncorrected). In each subject, the clear vs. degraded speech contrast showed bilateral superior temporal gyrus activation ($p < 0.001$, uncorrected). Less consistent was activation of inferior frontal and premotor cortex. In contrast, degraded stimuli preferentially activated insula, anterior cingulate, orbitofrontal cortex, and prefrontal cortex ($p < 0.001$, uncorrected). Strength of activation in these brain regions varied across the subjects in each contrast. Ongoing ECoG analyses will determine the oscillatory correlates of regions activated in the fMRI scan for the same subject. To the extent the ECoG coverage overlaps we hope to uncover the pattern of connectivity among the language network.

These results identify a network of high-level language processing regions at the individual subject level in pre-surgical epilepsy patients and support findings from previous studies. In addition to guiding neuroscientific investigations of language processing, we will be able to compare network level activation between fMRI and ECoG. As language localization is critically important for surgical planning in epilepsy patients, we plan to compare the results we obtain from the localizer task to the gold-standard epilepsy surgical planning data from the Wada test and electrical stimulation mapping results.

95. Top-down and bottom-up predictions in auditory decision-making

Lalitta Suriya-Arunroj, Yale E. Cohen, Joshua I. Gold

Auditory perceptual decision-making is mediated by both bottom-up (i.e., incoming sensory stimuli) and top-down (e.g., different cognitive and task-relevant pieces of information) processes. For example, changes in stimulus salience, such as regularity violation, can bias perceptual decision-making through bottom-up attention. Likewise, prior knowledge can also modulate decisions by prioritizing specific features of stimuli (top-down processing). However, the effects of top-down and bottom-up processing on auditory decision-making are not well understood. We investigated how both bottom-up regularity violations and top-down instructed expectations affect auditory decisions.

Human subjects listened, via headphones, to sequences of high-frequency (2000 Hz) or low-frequency (500 Hz) tone bursts (duration: 300 ms with a 10-ms cos² ramp; inter-burst interval: 100 ms) that were embedded in background of broadband noise. The subjects reported whether the last tone in each sequence (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 sound level of the test tone relative to the noisy background. 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 (5:1 and 1:5), and (2) the number of pre-tones, which were selected from an approximately exponential distribution (range: 0–14) to minimize the ability of the subject to predict the onset of the test tone. Top-down processing was manipulated by presenting, before sequence onset, a visual cue that indicated the prior probability that the test tone could be high or low frequency within a given block of trials (corresponding to ratios of low- versus high-frequency test tones of 5:1, 1:1, 1:5).

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. For the top-down manipulation, increasing the prior probability for one alternative biased choices and simultaneously decreased response times for that alternative. For the bottom-up manipulation, choices were biased toward the frequency value of the preceding tones, whereas choices toward the opposite frequency showed reduced reaction times. Further analyses and modeling of choices and response times in a drift-diffusion framework will help identify the computations used to incorporate these two sources of information into perceptual decisions.

96. The effect of anticipated cue reliability on behavioral and neural adaptation in barn owls

Michael V Beckert, Roland Ferger, Jose L Pena

The brain actively updates the representation of the environment. An open question about this function is whether adaptation is weighted by the predicted statistics of sensory information. Here we asked whether anticipated cue reliability affects the rate of adaptation in the auditory system of the barn owl.

The midbrain of the barn owl contains a map of auditory space, which uses the interaural phase difference to compute sound location in azimuth. Previous work showed that space-specific neurons in this map are tuned to the frequency range that is most reliable for its preferred location. This effect is due to the acoustical properties of the head, causing higher frequencies to convey interaural phase difference more reliably in frontal space and lower frequencies in the periphery. We hypothesized that adaptation to sound location would differ between stimuli expected to be reliable or unreliable. We tested this hypothesis at the behavioral and neural-population levels. We measured the pupillary dilation response, an orienting response that readily adapts upon repetition of a stimulus. Tones were repeatedly presented through earphones to an awake barn owl, either from the front or periphery of the head to measure the habituation rate. To assess the strength of the novelty detection, a deviant in location was then presented to elicit a recovery of the PDR. Trials with higher cue reliability were compared to trials with low reliability to find trends in habituation. To assess this question at the neural-population level, we conducted recordings of multiple neurons in the space map using a microelectrode array. Adapter and test stimuli were used to assess population and activity and tuning of individual cells before and after adaptation. Frontal and peripheral neurons were compared to test the hypothesis that anticipated reliable and unreliable stimuli lead to different adaptation rates.

97. Decoding of phonemes in continuous speech from fMRI response patterns

Julia Erb, Dino Düwel, Giancarlo Valente, Federico De Martino, Elia Formisano

Speech comprehension entails the extraction of linguistic features from the acoustic signal. Phonemes are thought to be the first level of abstraction from the acoustic information. How is phonemic information encoded in the cortex? Here, we used an open high-resolution 7T-fMRI data set by Hanke et al. (2014, 'study forrest') where humans listened to a continuous, naturalistic audio movie. We used multivariate regression to reveal the representation of German phoneme classes in the auditory cortex. Results indicate that fMRI responses to the temporal distribution of phoneme classes in continuous speech can be decoded with a mean accuracy of 65% ($p < 0.001$, permutation test). Maps of the most discriminative voxels revealed a distinct spatial distribution across tonotopic fields for different phoneme classes. Thus, we show the feasibility of multivariate linear regression to decode the phonemic information from fMRI responses to continuous, naturalistic speech.

98. The Role of Auditory Experience in the Neural Systems for Effortful Listening

Bradley E. White, Clifton Langdon

INTRODUCTION

Behavioral and neuroimaging research suggests that processing degraded acoustic information creates a cascading effect on the mechanisms underlying speech comprehension, suggesting that our cognitive resources are limited and causing a trade-off between effort and comprehension [1,2]. Here, using a plausibility judgment task and functional near-infrared spectroscopy (fNIRS), we aim to dissociate motivated listening and its modulation of language processing in response to increased demands on executive functioning in typically-hearing listeners and experienced, early-deafened hearing aid (HA) and cochlear implant (CI) users.

Hypotheses. We hypothesize that early-life chronic exposure to acoustically degraded speech will result in a neural network for spoken language processing that has adapted to degradations in the acoustic signal, thus increases in processing demands in HA and CI users will elicit neural networks for listening that are similar to their typical hearing counterparts. Alternatively, we hypothesize that early-life chronic exposure to acoustically degraded speech will result in a neural network for spoken language processing that reflects persistent contribution from cognitive attention and short-term verbal working memory mechanisms, thus everyday listening and increases in processing demands in HA and CI users will elicit dissimilar neural networks for listening that are dissimilar to their typical hearing counterparts.

METHODOLOGY

Participants. Three participant groups are recruited, all of which are monolingual English-speaking adult participants (ages 18;0 to 40;11.) (1) Typical hearing control participants have clinically-defined typical hearing, characterized by audiometric three-frequency pure-tone averages (i.e., 500 Hz, 1 kHz, and 2 kHz) ≤ 25 dB HL and no reported history of hearing loss. (2) HA users have moderate to profound hearing loss, characterized by pure-tone averages >40 dB HL. (3) CI users (both uni- and bilateral recipients are included.) Both experimental groups will have begun using HAs or CIs $\leq 5;0$ years old and demonstrate good speech perception capabilities. Participants in all 3 groups will be matched based on psychometric measures of intelligence and English language proficiency.

Procedures. Participants complete a battery of language and cognitive assessments. The fNIRS task presents participants with 288 sentences for a plausibility judgment task. The sentences vary linguistically (i.e., simple subject-relative and complex object-relative clause structures) and acoustically (i.e., regular speed and speeded). Control participants were also presented with other acoustic variations (i.e., clear speech, moderately distorted hearing aid simulation, and severely distorted 8 channel noise-vocoded speech).

RESULTS

Results will show a relationship between auditory experience and the neural systems for spoken language processing. The strength of this relationship will be assessed using group

comparisons (e.g., ANOVAs) of TH, HA, and CI listeners' hemodynamic responses while listening to syntactically and acoustically variable speech. These results will allow us to better understand the cognitive mechanisms that underlie "successful" HA and CI users.

CONCLUSION

Differentiating between these two hypotheses advances our understanding of how auditory experiences modulate language and associated processes (e.g., attention, short-term verbal working memory), and specifically the relationship of this modulation to the amount of speech features preserved in the signal.

REFERENCES

- [1] Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear and hearing*, 39(2), 204.
- [2] Alain, C., Du, Y., Bernstein, L. J., Barten, T., & Banai, K. (2018). Listening under difficult conditions: An activation likelihood estimation meta-analysis. *Human brain mapping*.

99. Cortical network topology across awareness states during sleep and anesthesia: An intracranial electrophysiology study

Matthew I. Banks*, Kirill V. Nourski*, Hiroto Kawasaki, Matthew A. Howard III

Introduction: The sharing of information between nodes in the cortical network plays a central role in leading theories of consciousness, and disruption in connectivity has been proposed to occur upon loss of consciousness (LOC) during anesthesia and sleep. However, whether LOC during these two conditions shares a common mechanism is unclear. To investigate this issue, resting state network topology was compared across brain states during natural sleep and propofol anesthesia.

Methods: Subjects were neurosurgical patients implanted with intracranial electrodes placed to identify epileptic foci. A combination of subdural grids and depth electrodes provided dense coverage of temporal, parietal and frontal cortex. We focused on nodes in the cortical hierarchy activated during both pre-attentive and conscious auditory novelty detection: core and non-core auditory cortex on the superior temporal gyrus including the superior temporal plane, auditory-related cortex on the middle temporal and supramarginal gyrus, and prefrontal cortex. Resting state data were recorded in the same subjects during overnight natural sleep and during induction of general anesthesia with incrementally titrated propofol infusion. Six brain states were compared: wake (WS) and NREM stages 1 and 2 (N1, N2) during natural sleep, and pre-drug wake (WA), sedated/responsive (S) and unresponsive (U) during propofol anesthesia. Adjacency matrices (A), computed as thresholded, weighted, alpha (8-13 Hz) phase lag index, were compared pairwise for brain states using the operator norm of the difference between adjacency matrices (i.e. $d_{i-k} = ||A_i - A_k||_{op}$).

Results: Changes in network topology were more dramatic for transitions into the unconscious states (N2, U) than for transitions into states of diminished but maintained awareness (S, N1) (i.e., $d_{WA,S} < d_{S,U}$ and $d_{WS,N1} < d_{N1,N2}$). Network topology was most similar between brain states hypothesized to be equivalent under sleep and anesthesia (i.e. WA vs. WS, S vs. N1, U vs. N2); d values comparing hypothesized equivalent states (e.g. $d_{U,N2}$) were smaller than d values for corresponding non-equivalent states (e.g. $d_{S,N2}$).

Conclusions: Pronounced changes in network topology for the transitions S → U and N1 → N2 likely reflect changes in cortical connectivity mediating transition between conscious and unconscious states. The similarity in network topology between equivalent brain states during anesthesia and sleep suggests common mechanisms in transitions to and from unconsciousness.

100. Triadic forebrain structures that directly control the auditory midbrain of echolocating bats

T. Ito, R. Yamamoto, T. Furuyama, K. Hase, K. I. Kobayashi, S. Hiryu

Echolocation bats change spectrotemporal parameters of sonar pulses during navigation. Furthermore, echoes may convey emotional values which interfere or modify the plan of pathways. Since the echoes arrive several milliseconds after the emission of pulses, the inferior colliculus (IC), the midbrain auditory nucleus that analyzes the echoes, must be quickly adjusted for each pulse. Here, we demonstrate that three forebrain structures, namely infralimbic cortex (IL), magnocellular part of the basal nucleus of the amygdala (Bmg), and auditory cortex (AC), send direct descending projection to the IC by using a retrograde tract tracing method. All three structures projected to bilateral IC although the ipsilateral projection was dominant. Comparisons of pattern of retrogradely labeled cells across animals suggested that ipsilateral AC projection to the IC is tonotopically organized. Projections from other forebrain structures did not show clear tonotopicity. Together with evidence of previous studies, these results demonstrated the triadic descending projections to the IC which make loops between forebrain and IC. As IL, Bmg, and AC relate to navigation, emotional value, and spatial coding, respectively, the loops may quickly optimize active sensation during navigation.

101. Electrocorticographic (ECoG) analysis of dialog-based paradigms for assessing speech, language and cognitive functions: a case report

Mitchell Steinschneider, Kirill V Nourski

Acquisition of ECoG data while subject/patients engage in a dialog-based study permits high resolution analysis of multiple speech, language and cognitive functions in a concise question/answer format. Analysis focuses on neural processing associated with listening, performing mental calculations, and verbal responses. Here, we demonstrate the utility of this paradigm by presenting data from a single neurosurgical subject with normal hearing and cognitive functions, and who had extensive right hemisphere electrode coverage of temporal, temporo-parietal, and frontal cortices. The paradigm was an expanded version of the Mini-mental State Examination, which included additional spelling, naming, and memory-based tasks. Recording of the verbal exchange was parsed using Praat software based upon natural articulatory breaks in the conversation. Cortical recording sites were categorized based upon their location, within anatomically defined regions of interest (ROIs), including e.g., Heschl's gyrus, posterior, middle and anterior portions of the lateral superior, middle and inferior temporal gyri (STG, MTG). A key analysis was the degree of high gamma (70-150 Hz) activation during listening to the interviewer vs. during one's own speech (Nourski, Steinschneider, Rhone, *Front Hum Neurosci* 2016 10:202). ROIs where listening was associated with stronger high gamma activity than speaking were restricted to the posterior and middle portions of the STG. ROIs where speaking was associated with stronger high gamma activity than listening included Heschl's gyrus, planum polare, anterior STG, middle/anterior MTG, ITG, supramarginal gyrus and the temporal pole. A second analysis examined the relationship between high gamma and activity in lower ECoG frequency bands. In all ROIs, there was a positive correlation between high and low gamma (30-70 Hz). Relationships between high gamma and alpha (8-14 Hz) and theta (4-8 Hz) bands were inconsistent across ROIs. Largest high gamma responses were generally associated with relatively difficult tasks, naming favorite items and task completion. We conclude that monitoring one's own speech can extend beyond the classically defined dorsal auditory-motor speech pathway into the ventral pathway involved in listening and decoding speech at progressively higher processing levels. A consistent relationship between high and low gamma activity supports the utility of low gamma acquired in non-invasive studies as a proxy for high gamma activity. Comparing neural activity across subjects may assist in defining the natural variability in language and cognitive processing strategies utilized by individuals.

102. Memory and integration of faces and vocalizations in neuronal populations in the primate prefrontal cortex

Shraddha Shah, Theodore Lincoln, Katryna Kevelson, Lizabeth M. Romanski

The ventrolateral prefrontal cortex (VLPFC) plays an important role in language and communication. Previous studies have shown that VLPFC neurons respond to and integrate face and vocalization stimuli and hold information about these stimuli online during working memory. Prefrontal neurons are highly selective in their responses to specific auditory and visual stimuli. Moreover, the magnitude and direction of multisensory responses is dependent upon particular combinations of faces and vocalizations, as well as their temporal synchrony and semantic congruence. Furthermore, neighbouring cells show responses to similar modalities and categories of stimuli. This suggests that there may be a topographic organization of the features encoded and integrated by single prefrontal neurons. However, real world approximations of brain function require that we investigate memory and perception at the level of cell ensembles. Thus, in the current study, we used multiple single electrodes and linear electrode arrays to record simultaneously from multiple neurons while nonhuman primates performed an audiovisual working memory task using dynamic face and vocalization stimuli as the memoranda. Responsive neurons were further tested with additional exemplars of faces, vocalizations and their combination. We manipulated the stimuli along relevant feature axes including identity, vocalization category, semantic congruence, and temporal synchrony to investigate the salience of these features. We recorded from pairs and clusters of prefrontal neurons and assessed stimulus selectivity, task parameter selectivity and functional interactions between neurons within and across ensembles. By investigating the multisensory response structure at the level of neuronal ensembles in this manner, we will be able to determine what aspects of communication stimuli are being encoded and integrated by prefrontal neurons.

103. Are human hippocampal sharp wave ripples involved in auditory sequence encoding?

Araceli Ramirez-Cardenas, Juan Felipe Ramirez-Villegas, Matthew Howard, Jeremy Greenlee

Sharp-wave ripples (SWRs) are brief high-frequency oscillatory field potentials generated in the hippocampus and involved in memory consolidation. These events are typically described in mammals during quiescent and offline states (slow-wave sleep, immobility), when neural activity reflecting recent experience is replayed. Hippocampal replay during sharp-wave ripple events (SWRs) is thought to drive neocortical plasticity and memory consolidation. However, SWRs have been recently reported during active waking states in rats and primates. For example, they appear during goal-directed visual exploration in macaques (Leonard et al. 2015, Leonard & Hoffman, 2017). The rate of these exploratory hippocampal SWRs increased with repeated presentations, while their likelihood increased near the remembered object. These results could suggest that SWRs encode repetition when an active search is driven by past experience.

SWRs seem to also play an important role in hippocampal-cortical communication underlying memory consolidation. Indeed, SWRs have been shown to drive excitatory responses in output cortical areas, including the auditory cortex. In the context of a sound-guided task a recent animal study showed that during sleep there is a rapid cortical-hippocampal-cortical loop of information flow driving the consolidation of auditory memories (Rothschild, et al. 2016). Here we performed electrocorticography recordings during a task involving the presentation of auditory sequences. Our preliminary analysis look at the role of hippocampal sharp-wave ripples in the learning of auditory sequences repetitively presented both passively and in the context of a goal-directed task. We further look at the role of SWRs in the interactions between the auditory cortex and hippocampus.

104. Behavioral and electrophysiological evidence of incidental learning across continuous speech

Yunan "Charles" Wu, Ran Liu, Sung-Joo Lim, Lori Holt

Language learning requires that listeners generalize across acoustically variable speech sounds to recognize linguistically-relevant units like words and phonemes. Although a growing literature on auditory category learning guides understanding of the mechanisms available to support this learning, most studies have examined learning across isolated sound exemplars like individual phonemes or words. In contrast, real-world learning typically takes place across fluent, continuous speech. How do listeners learn across acoustically variable continuous speech when they do not have a priori information about the temporal window across which learning must occur? We hypothesized listeners discover temporal 'islands of reliability' in highly variable, continuous speech signals that are consistently associated with behaviorally relevant actions and events. Further, we predicted that listeners learn to treat the acoustics in these temporal windows as functionally equivalent, leading to behavioral and representational change consistent with category learning. Here, N native-English participants played a videogame in which actions directed at alien creatures were consistently associated with acoustically-variable Mandarin Chinese target words embedded in continuous Mandarin speech spoken by 4 native talkers (2 female). Control words were also embedded in the continuous speech and were presented equally as often as target words; but control words were not associated with any particular alien or event. Participants played the videogame for 3.5 hours across 5 days, with no prior knowledge of Mandarin. Neither overt categorization decisions nor overt feedback were involved in the videogame. Following training, an overt post-training categorization test revealed robust learning of target words that persisted at least 10 days and generalized to novel utterances and talkers. Comparison of pre- versus post-training electroencephalography responses to continuous Mandarin revealed that target words, but not control words, elicited an enhanced auditory evoked N100 response in frontal and central electrodes associated in prior research with word segmentation. This suggests that incidental learning under conditions in which behaviorally relevant actions and events align functional units in continuous speech leads to more robust speech category learning than passive distributional learning through mere exposure.

105. Quadratic discriminant analysis reveals representational dissimilarities between different types of auditory selective attention

Winko W. An, Alexander Pei, Barbara G. Shinn-Cunningham

Introduction

Representational similarity analysis (RSA) has been proven successful in integrating MEG with fMRI data through the use of a representational dissimilarity matrix (RDM). Previous studies used support vector machine (SVM) to build MEG RDMs that could effectively differentiate conditions of different categories. How to build optimal RDMs for EEG studies, however, is yet to be explored. This study examined the viability of using quadratic discriminant analysis (QDA) to improve the decoding accuracy in an EEG RDM for a study of auditory selective attention.

Method

12 adults participated in this study. The auditory stimuli were syllables Ba/Da/Ga spoken by 5 distinguishable talkers, each played from one of 5 possible directions (left/right 90°/30°, center). The experiment consisted of 21 conditions (36 trials each). Trials differed in the type of attention being tested (spatial/talker/no attention) and the gender/direction of the target. For each trial, the subjects were given a visual cue, an auditory cue (cueing period), and a 1-second gap (preparatory period) before the onset of a 4-syllable mixture. The cues conveyed information about the type of attention to be tested, and the direction/talker of the target. For a spatial- or a talker-attention trial, the subjects were asked to report the syllable from the target direction/talker in the upcoming mixture. In control trials, subjects were asked to respond with a random button press, ignoring the stimuli. 64-channel EEG, recorded throughout the experiment, was pre-processed using a customized script. Each time point in the EEG data during the cueing and the preparatory period were used to train and test SVM/QDA for neural decoding. Results of SVM and QDA classifiers were compared in terms of the decoding accuracy during the two periods, i.e. the difference in evoked and induced responses, respectively. Student t-test was used for statistics.

Results

Both classifiers could decode Spatial vs Control (SpaCtr), Talker vs Control (TlkCtr) and Spatial vs Talker (SpaTlk) better than chance ($p < 0.01$) in both periods. QDA yielded higher decoding accuracy than SVM for the cueing period in SpaCtr (57.91%±1.14% vs 55.36%±0.76%, $p < 0.001$), TlkCtr (57.45%±1.47% vs 56.34%±1.23%, $p = 0.027$), and SpaTlk (58.06%±1.07% vs 55.47%±0.65%, $p < 0.001$). In the preparatory period, QDA outperformed SVM in SpaCtr (52.48%±1.04% vs 50.62%±0.65%, $p < 0.001$), TlkCtr (51.78%±0.49% vs 50.76%±0.50%, $p < 0.001$), and SpaTlk (51.97%±0.77% vs 50.51%±0.45%, $p < 0.001$).

Conclusion

QDA could reveal representational dissimilarities more effectively than SVM in EEG data, especially for induced responses.

106. Investigating audiovisual integration in the inferior colliculus

Aaron Williams, Katherine Wood, Melanie Tobin, Maria Geffen

Interaction between the auditory and visual systems is an important component of sensory processing, improving object localization and speech recognition, and facilitating the generation of a unified sensory percept of objects in the external environment. Audiovisual integration is largely studied in the cortex, however the contribution of subcortical areas to this process is poorly understood. The inferior colliculus is a primarily auditory region in the dorsal midbrain that receives robust cross-sensory input from modalities including the visual system. We seek to investigate audiovisual integration in this region, specifically the ability of visual input to modulate auditory responses. We hypothesize that temporal coherence between visual and auditory stimuli will enhance neural encoding of the auditory stimulus. Using calcium imaging, we have found that up to 30% of neurons in the external shell of the inferior colliculus are tuned to specific frequencies of auditory input. We implemented mutual information to objectively quantify and compare stimulus feature encoding by individual neurons, confirming that frequency-tuned neurons maintain higher mutual information with the stimulus compared to their non-tuned counterparts. We will build on this research by presenting visual and audiovisual stimuli to characterize multimodal tuning properties of the region. Additionally, we will vary the temporal coherence between the two stimulus streams to clarify whether the region is capable of binding stimulus features to improve sensitivity. Our studies will potentially reveal a novel role of the inferior colliculus in integrating these two sensory modalities, providing a foundation for further research on sensory integration and processing disorders.

107. Indexing multisensory integration of natural speech using canonical correlation

Aisling E. O'Sullivan, Michael J. Crosse, Giovanni M. Di Liberto, Joe Majeski, Alain de Cheveigné and Edmund C. Lalor

Speech is most commonly perceived as multisensory. Indeed, integrating auditory and visual information from a talker's face is known to benefit speech comprehension. However, the neural mechanisms underlying this integration are not well understood, especially in the context of natural, continuous speech. Recent work relating EEG to the acoustic speech envelope has shown enhanced tracking of congruent audiovisual (AV) speech relative to unisensory (A+V) speech (Crosse et al., 2015), especially under challenging listening conditions (Crosse et al., 2016). This approach of relating EEG activity to the acoustic envelope however, is limited in its ability to deal with more complex representations of the speech signal. This is unfortunate given recent work demonstrating the ability to index auditory speech encoding at different hierarchical levels using EEG (Di Liberto et al., 2015). In order to overcome this limitation we have used canonical correlation analysis (CCA) to relate a multivariate representation of a speech stimulus to the multivariate EEG response. Specifically, CCA applies a linear transformation to both the stimulus and response with the goal of optimizing the correlation between the two. This allows us to examine integration effects at different hierarchical levels using spectrotemporal and phonetic feature representations of speech. Our results show that when we represent the speech in terms of its spectrotemporal information there is a significant multisensory integration effect for speech in noise – suggesting that seeing the speakers face restores tracking of the spectrotemporal information in the speech signal. When we represent the speech in terms of its phonetic content however, we find a significant multisensory effect both for speech in quiet and speech in noise. Thus it appears that having access to the visual articulations of the speaker benefits phonetic encoding of the speech signal above what the acoustic information alone can provide. Further analyses will seek to isolate the unique contributions of spectrotemporal and phonetic processing to the EEG signal. The overarching goal is to provide a framework for testing hypotheses about how the temporal dynamics and articulatory information from a speaker's face help us to understand speech in challenging listening conditions.

108. **MRI monitoring of macaque monkeys for neuroscience: Two case studies**

Fabien Balezeau*, Jennifer Nacef*, Kathy Murphy, Yukiko Kikuchi, Felix Schneider, Francesca Rocchi, Ross S. Muers, Rocio Fernandez-Palacios O'Connor, Christoph Blau, Richard Saunders, Matthew Howard III, Alexander Thiele, Timothy D. Griffiths & Christopher I. Petkov

Information from Magnetic Resonance Imaging (MRI) can be useful for managing infrequent neurological or neuroscientific cases, provided that there are no MRI contra-indications precluding scanning. Here we report on the utility of MRI-based monitoring in two cases with rhesus macaques. In both cases the monkeys presented with subtle to mild clinical signs, were well otherwise and without a significant increase in welfare impact, hence they were identified as suitable candidates for clinical investigation, MRI-based monitoring and treatment. The first case (M1) presented with left-handed weakness contralateral to a recording chamber over sensory motor cortex above the central sulcus. T1 and T2 weighted MRI imaging (4.7 Tesla, Bruker) identified two suspected sub-dural abscess sites. These were targeted for antibiotic treatment and either aspiration or implant modification. Thereafter the animal's hand use returned to normal. The second case (M2) presented with left-handed weakness from a different basis. T1, T2 and proton density scans identified a suspected internal cerebral haemorrhage in the acute/subacute stage (hypo-intense signal on both T1 and T2), presenting as an area of MRI signal loss. Continued MRI monitoring coincided with a return to unremarkable behaviour and a substantial reduction in the size of the affected area combined with signal resolution under which normal tissue could once again be observed. Time of flight MRI angiography identified a subdural venous drainage network and recordings were planned to avoid it. In summary, MRI assists in a precise diagnosis of cerebral events and can be a valuable addition to clinical treatment to ensure resolution.