8th International Conference on Auditory Cortex

Abstract booklet

Maastricht September 14th-18th.

Crowne Plaza Hotel

Ruiterij 1, 6221 EW Maastricht, The Netherlands

Index

Invited Talks

ID: 109

Invited talks

Keywords: Perception, stress, auditory cortex, 2-photon

Repeated stress gradually impairs auditory processing and perception

Jennifer Resnik

Ben-Gurion University, Israel; resnikj@bgu.ac.il

Repetitive stress is a prevalent aspect of modern life and a significant risk factor for various psychiatric and sensory disorders. While perceptual abnormalities are common in these conditions, the specific impact of repetitive stress on sensory processing and perception remains underexplored. In this talk, I will present our findings on how repetitive stress reshapes auditory cortical activity and alters perceptual processing in adult mice. Utilizing a combination of chronic stress paradigms, longitudinal cortical activity measurements, and auditory-guided behavioral assessments, we observed that repetitive stress leads to increased spontaneous cortical activity and diminished sound-evoked responses in both pyramidal and parvalbumin-expressing neurons. Conversely, somatostatin-expressing neurons exhibited heightened sound-evoked responses under stress. These neural alterations were associated with perceptual changes, notably a reduction in loudness perception. I will discuss how these findings provide insight into stress-induced plasticity, potential underlying mechanisms, and broader implications for sensory dysfunctions in psychiatric disorders.

ID: 125
Invited talks
Keywords: auditory cortex, projection neurons, decision making

Brain-wide Neural Circuits for Sensory-Guided Behavior

Ross S. Williamson
University of Pittsburgh, PA, USA; rsw@pitt.edu

Auditory-guided behavior is ubiquitous in everyday life, whenever auditory information is used to guide the decisions we make and the actions we take. One such behavior is auditory categorization, a process that reflects the ability to transform bottom-up sensory stimuli into discrete perceptual categories and use these perceptual categories to drive a subsequent action. Although this process is well-documented at the behavioral and cognitive levels, surprisingly little is known about the explicit neural circuit mechanisms that underlie categorical computation and how the result of this computation drives behavioral outcomes. We believe that the transformation of auditory information into an appropriate behavioral response is necessarily a brain-wide endeavor. The deep layers of the auditory cortex give rise to several massive projection systems that exert influence over many downstream brain areas. Here, I will discuss our efforts towards understanding the organization of these projection systems and how they differentially contribute to auditory-guided behavior.

ID: 143
Invited talks
Keywords: pitch, ferret, aging, perception, auditory cortex

Cortical representations of pitch for effective real-world listening

Kerry Marie May Walker
University of Oxford, United Kingdom; kerry.walker@dpag.ox.ac.uk

Pitch is our perception of the tonal quality of sounds, which corresponds to the sound's fundamental frequency (F0). Pitch is a key feature of our everyday experience of speech and music, as well as our ability to attend to a single voice in a crowded room. Lesion evidence suggests that auditory cortical activity plays an essential role in pitch perception, but the cellular mechanisms through which F0 is derived from basic acoustical cues such as resolved harmonics (frequency-based cues) and periodicity (temporal cues) remain poorly understood. I will discuss our Neuropixels recordings in ferrets, which examine how auditory cortical neurons combine basic acoustical cues to derive a robust representation of pitch that is invariant to changes in sound source. Our data suggest that this invariant pitch encoding, which would allow us to recognize a melody across different instruments, is not trivially predictable from a neuron's pure tone frequency tuning. We find that such invariant representations of pitch are present in a small subset of individual auditory cortical neurons (putative "pitch neurons"), but can also be decoded from the population responses of less specialized neural ensembles. Furthermore, our psychophysical experiments show how ferrets and human listeners use pitch as a streaming cue to attend to a target sound source while ignoring others. We show how reliance on periodicity and resolved harmonic cues for selectively attending to a voice changes across the human lifespan and in hearing loss. Together, these studies help us understand how the brain derives our percept of pitch from resolved harmonics and temporal periodicity. The results may help improve speech-in-noise intelligibility for people with hearing impairments.

ID: 148
Invited talks
Keywords: pitch perception, cross-cultural

Pitch representations are richer and more diverse than previously thought

Malinda Jeanette McPherson-McNato
Purdue University, United States of America; mjmcp@purdue.edu

Pitch is a defining property of the most important sounds humans hear, including speech and music. Acoustically, sounds said to have pitch have a regular rate of repetition in time – a “fundamental frequency” or f0 – and contain overtones (‘harmonics’) that are multiples of the f0. Pitch is traditionally construed as the perceptual correlate of f0, and a longstanding goal of hearing research has been to determine how listeners estimate f0 from harmonic sounds. However, recent work suggests that this classic view of pitch as f0 estimation is incomplete – at least two representations are involved, one of which does not involve f0. In this talk, I will present evidence demonstrating that listeners can often estimate pitch changes by tracking the frequency spectrum without estimating f0. Nonetheless, representations of harmonic structure and the f0 appear to help listeners compress sounds into compact representations that aid memory. These compact representations may be critical for music perception, where remembering the pitches of notes across longer time spans is necessary. I will discuss recent cross-cultural work demonstrating that this basic structure of pitch perception is shared across cultures despite other differences in pitch-related behavior. I will also discuss fMRI results examining how these pitch representations are encoded in the auditory cortex. Overall, this work demonstrates that pitch perception is mediated by several different mechanisms, only some of which conform to traditional notions of pitch.

ID: 151
Invited talks
Keywords: Prediction, memory, electrophysiology, neuroimaging, modelling

Auditory prediction and memory in context

Ryszard Auksztulewicz
Maastricht University, Netherlands, The;
ryszard.auksztulewicz@maastrichtuniversity.nl

The brain is believed to generate internal predictions based on previous experiences to optimize behavior, a phenomenon known as predictive processing. This process has been extensively validated in both human studies and animal models, yet the extent to which contextual factors, such as attention and task demands, modulate auditory predictions remains inadequately characterized. In this presentation, I will elucidate findings from a series of investigations that integrate neural recordings from both humans and rodents with computational modeling to delineate the brain mechanisms underlying sensory predictions and their interactions with cognitive processes.

Our electrophysiological and neuroimaging studies conducted in humans indicate that the effects of predictions are context-dependent rather than automatic, with behavioral and neural data analyses revealing that the relevance of predictions modulates their influence. Generative computational modeling links different types of prediction - e.g., for stimulus contents vs. timing - to dissociable mechanisms, including different forms of gain control in sensory processing regions. Concurrently, studies employing direct neural recordings from anesthetized rodents demonstrate that auditory cortical activity can decode neural representations associated with stimulus memory and predictions, suggesting an evolutionary conservation of predictive processing mechanisms across species.

In more recent work, an electroencephalography study investigating predictions during associative learning reveals that the dynamics of stimulus predictability are dynamic at both shorter (peri-stimulus) and longer (across learning) time scales. Specifically, valid predictions of visual stimulus categories enhance decoding at short latencies, indicative of representational sharpening, while they lead to reduced decoding at longer latencies, consistent with suppressive effects. Notably, across multiple trials, early stages of associative learning exhibit putative dampening effects, whereas representational sharpening requires a longer temporal buildup. These effects are further linked to region- and layer-specific effects using high-field neuroimaging.

Collectively, these findings emphasize that predictive processing is governed by context-specific neural mechanisms rather than a homogeneous process. This research underscores that auditory perception is a highly dynamic process under modulation by cognitive and contextual factors, integrating past experiences with current stimuli and future predictions.

ID: 157
Invited talks
Keywords: GCaMP8m, Prediction, Auditory, Mouse, Calcium imaging, CBA

The representation of omitted sounds in the mouse auditory cortex

Bernhard Englitz
Radboud University, Netherlands, The; bernhard.englitz@donders.ru.nl

Humans and animals use predictions from past experience to optimize their behavior, however, its neuronal implementation has remained elusive. We address this implementation on the macro- and microscale of the mouse auditory cortex using high-speed imaging.

Neural responses to the omission of expected sounds were precisely timed, localized to a higher auditory area (TeA) and continued to rise until the following stimulus. This spatial representation was distinct from that of offset responses, deviant sounds and stimulus-specific adaptation, and most prominent in L1-4. Omission responses were also temporally distinct from offset responses. Animals were aware of omissions and sequence statistics, as indicated by timed pupil dilation and rapid facial motions.

The localized omission response in TeA supports a hierarchical organization of predictive processing. However, the continued rise suggests an integrated prediction error, instead of a representation of prediction or prediction error, expected to terminate with the omission period by predictive coding.

ID: 239
Invited talks
Keywords: Discrimination in noise, behavioral performance, neural correlates, inferior colliculus, inactivation

Is auditory cortex necessary for the discrimination of communication sounds in noise?

Jean-Marc EDELINE, Samira SOUFFI, Alexandra MARTIN, Chloé HUETZ
NeuroPSI CNRS & Univ Paris-Saclay, France; jean-marc.edeline@universite-paris-saclay.fr

In the auditory system, as in the other sensory modalities, the hierarchical organization – the higher in the system, the more elaborated is the processing - is still the dominant view, which largely impacts on the results interpretation and on the papers acceptance in high impact factor journals. The consequences of this dogma are that the most difficult and "cognitive" is considered to be the task in which you engage an animal, the more we suspect that the cortex is necessary for a good performance in this task. Among many other functional roles, recent papers suggest that the auditory cortex is promoting (i) auditory stream segregation, (ii) egocentric and allocentric representations (iii) pitch judgment (iv) latent knowledge and (v) tinnitus perception. In fact, most of these claims were often made based upon correlations between cortical responses and behavioral performance without testing whether or not similar correlations could have been observed in subcortical auditory structures. In a set of studies, our team has compared the neuronal discrimination performance of cortical and subcortical neurons at presentation of four conspecific (or heterospecific) vocalizations presented in quiet then in various levels of noise. Both in guinea pig and mice, the inferior colliculus neurons showed higher discrimination abilities than the thalamic and cortical neurons, and their responses were more robust to noise than the cortical and thalamic ones. To go further, we assessed whether behavioral performance indeed relies on cortical or subcortical activity by pharmacologically silencing the auditory cortex during training. In several mice, discrimination performance was intact both in quiet and in noise (up to -10dB SNR) during cortical silencing.

These results indicate that very good performance levels in discriminating between vocalizations can be achieved based only on the activity of subcortical structures, even in situations of acoustic degradations. We cannot rule out the possibility that the cortex is required during the initial steps of the discrimination task, but it seems that once discrimination is mastered (i.e. after an extensive overtraining) excellent perceptive performance (including in difficult noisy conditions) can be achieved based on the processing performed by subcortical neurons. Together, our data fit with the Reverse Hierarchy Theory proposed in the visual system, which asserts that learning is mainly a top-down process, beginning at high level areas and which progresses backward to the input levels where the discrimination abilities are optimal.

ID: 244
Invited talks
Keywords: auditory attention, medial geniculate body, ultra-high field MRI, computational modelling

Probing the role of the human corticothalamic circuitry in auditory attention

Michelle Moerel[1,2]
[1]Maastricht Centre for Systems Biology and Bioinformatics (MaCSBio), Faculty of Science and Engineering, Maastricht University, the Netherlands; [2]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, the Netherlands; michelle.moerel@maastrichtuniversity.nl

Hearing is remarkably flexible. We can effortlessly focus on specific sounds, like an engaging conversation, amidst a cacophony of background noise. This ability is thought to rely on dynamic changes that rapidly optimize auditory processing for the task at hand. While higher-order cortical regions play a major role by directing attentional resources, recent animal studies highlight a crucial role for signals travelling between auditory thalamus and cortex, the corticothalamic circuitry, in auditory attention. Studying this circuitry in human brain is challenging due to the limited spatial resolution of non-invasive neuroimaging techniques. I will show how this challenge can be addressed by combining ultra-high field (UHF) MRI at 7 Tesla with computational modelling. I will show evidence of dynamic changes in auditory cortical brain processing that may underlie our selective listening capabilities, alongside preliminary data from our efforts to investigate the role of the human corticothalamic circuitry in auditory attention. Furthermore, I will discuss the relevance of these explorations for studying tinnitus, a condition often linked to disruptions in the circuitry between auditory thalamus and cortex.

ID: 251
Invited talks
Keywords: primate models, vocalisation, oscillations, neuropsychiatric disorders, neural circuits

Targeted neuromodulation of fronto-temporal circuits via auditory stimulation in nonhuman primate models

Yukiko Kikuchi
Biosciences Institute, Newcastle University, United Kingdom;
yukiko.kikuchi@newcastle.ac.uk

Disruptions in fronto-temporal networks are increasingly recognised as key contributors to the pathophysiology of various neuropsychiatric disorders. Oscillatory activity within these circuits is critical for cognitive functions such as memory and emotional regulation. However, current non-invasive interventions lack the spatial and temporal precision required to effectively target these networks for therapeutic benefit. This underscores the need for translational animal models to improve the efficacy of neuromodulation strategies. We investigate the causal role of auditory processing in modulating memory and emotion circuits using nonhuman primate models with the goal of developing novel auditory-based neuromodulatory interventions for neuropsychiatric disorders. Our recent genetic and behavioural work supports that nonhuman primates are ideal models for studying core symptoms of social cognitive disorders (Kikuchi et al., 2025), which are hypothesised to involve impairments in prediction. In prior studies, we demonstrated that during predictive sequence learning with speech sounds, prediction and prediction error signals are dissociable from stimulus-driven effects, exhibiting prolonged latencies manifested in theta-gamma coupling within the macaque and human auditory cortex. These findings suggest the presence of feedback prediction signals originating from hierarchically higher-order brain regions beyond the auditory cortex. Using multi-channel laminar electrophysiology and optogenetics, we extended this framework to auditory-frontal circuits, demonstrating that while the auditory cortex encodes individual sound identities, the frontal cortex tracks sequence ordering over longer timescales (Errington et al., ICAC 2025). In primates, hierarchically organised auditory ventral streams exhibit increasing selectivity and latency to complex sounds such as vocalisations, progressing along the caudorostral axis to the temporal pole and rostral superior temporal gyrus – regions directly connected to frontal and limbic circuits involved in emotional regulation and memory. We are developing closed-loop auditory stimulation to synchronise emotionally salient sounds in real-time with upstream endogenous oscillations of these circuits in marmosets and humans. These tools allow for measurement of physiological and behavioural stress markers linked to oscillatory modulation strength in freely moving animals (Woolgar et al., ICAC, 2025). This approach enables precise, causal analysis of circuit-level dynamics and their relationship to behavioural phenotypes relevant to psychiatric conditions, with the goal of defining circuit-specific, auditory-based interventions that can inform future therapeutic strategies. Supported by MRC(MR/X003701/1), BBSRC(BB/Z00005X/1) and EPSRC-CloseNIT (EP/W035081/1).

ID: 262
Invited talks
Keywords: temporal integration, internal state, sleep, ferret

Temporal integration in the auditory cortex across across cortical areas and vigilance levels

Hortense Gouyette[1], Magdalena Sabat[1], Sam Norman-Haigneré[2], Yves Boubenec[1]
[1]École Normale Supérieure, France; [2]Rochester University, USA; boubenec@ens.fr

A central question in auditory neuroscience is how the brain integrates sensory information over time to support perception of natural sounds. Natural soundscapes, including speech and vocalizations, contain structure spanning tens to hundreds of milliseconds, and must be parsed across multiple timescales. Yet it remains unclear whether this multiscale integration is achieved by heterogeneous populations of neurons, each tuned to a specific temporal window, or by individual units that dynamically adapt their integration depending on context, such as changes in sound statistics or internal state.

To address this question, we recorded neural activity using microelectrode arrays implanted across primary and non-primary auditory cortical fields in awake ferrets. By measuring the stimulus duration for which the neuronal response was invariant to surrounding context stimuli, we estimated the temporal window over which auditory cortex integrates incoming information. We found that neuronal responses were largely insensitive to sound structure occurring beyond a fixed, site-specific "integration window." These windows ranged from ~15 ms to over 150 ms and were hierarchically distributed, increasing systematically from primary to non-primary fields and across cortical layers. Importantly, the size of these integration windows did not vary with the information rate of the stimulus, suggesting that multiscale integration arises from a division of labor across neuronal populations, rather than dynamic adjustment by individual units.

We then investigated how temporal integration interacts with fluctuations in internal state. Physiological signals including breathing, heart rate, pupil diameter, and facial motion were monitored simultaneously. We observed that while these bodily variables modulated response gain and overall activity patterns in distinct ways, they had no measurable effect on the integration window itself. In contrast, pharmacologically induced sedation produced a consistent and significant lengthening of integration timescales. Preliminary data will investigate whether similar changes also occur during natural sleep states.

Together, these results reveal that the auditory cortex implements multiscale integration through a stable anatomical and functional hierarchy. While cortical responses are continuously shaped by the animal's internal state, the temporal window over which information is integrated remains remarkably fixed across spontaneous variations in arousal. Only major state transitions, such as sedation, appear to shift these timescales. These findings suggest that temporal integration is a robust and hardwired feature of cortical organization, essential for parsing the temporal complexity of natural sounds.

ID: 298
Invited talks
Keywords: Natural sounds, acoustics, semantics, convolutional neural networks

The perceptual and cerebral representation of natural sounds

Bruno Giordano
CNRS/Aix Marseille University; bruno.giordano@univ-amu.fr

Every day, we encounter diverse natural sounds (chirps, impulses) and effortlessly recognize the objects and events producing them (a bird call, a nail being hammered). Studying how we perceive these sounds provides a unique window into how the auditory system operates under real-world conditions, moving beyond the traditional focus on speech and music. Investigating these signals promises a broader understanding of auditory processing, encompassing both acoustic complexity and semantic structure.

Classical neuroscience work on natural sounds often focused on broad semantic classes (e.g., voice, music, animal calls), successfully identifying category-sensitive cortices but overlooking much of real-world semantic richness. Meanwhile, psychophysical research on everyday sound sources, such as how listeners discern material properties (wood, metal, glass) or actions (impact, scraping, rolling), has yielded new perceptual insights. Gaver's seminal taxonomy of everyday sounds offers a systematic approach for capturing these object- and action-related attributes, yet only a handful of neuroscience studies have integrated these ecologically informed perspectives.

Building on these foundations, my early work employed behavioral methods to examine how humans perceive impacted materials and the events behind them, culminating in a study showing how acoustic and semantic factors together shape the perceived dissimilarity of diverse natural sounds. This led to a first neuroimaging (fMRI) investigation demonstrating that the auditory cortex encodes both low-level acoustics and abstract object categories. More recently, collaborative efforts have combined behavioral measures, ultra-high-field (7T) fMRI, and convolutional neural networks (CNNs) to reveal that CNN-based models closely mirror perceptual and fMRI representations. Notably, mid-layer CNN representations best predict brain and behavioral data, likely emphasizing the acoustic structure critical for identifying objects and actions. In this sense, these intermediate representations could bridge natural sounds and speech processing through two distinct taxonomies: Gaver's and that of the International Phonetic Association.

Our ongoing work refines these approaches by optimizing stimulus selection, developing novel CNN models of the acoustic-to-semantic transformation, and tracking the temporal unfolding of cerebral representations through magnetoencephalography. Looking ahead, we aim to extend this framework from isolated sounds to the complexity of real-life auditory scenes, such as those found in offices or at the beach. By embracing the full richness of everyday listening, we seek a unified perspective on how the auditory system decodes our ever-changing acoustic environment, ultimately integrating neural, behavioral, and computational insights to illuminate the complexity of real-world hearing.

ID: 308

Invited talks

Keywords: learning, attention, plasticity, expertise

Active interaction shapes auditory perception and representation

Fred Dick

UCL, United Kingdom; f.dick@ucl.ac.uk

It is tempting to consider the auditory system as a kind of acoustic sea sponge, where sound information wafts through various filters that extract and process useful content while allowing irrelevant material to pass through. But auditory processing in humans and other mammals instead seems to be remarkably dynamic and responsive. It is shaped by our short- and long-term experience with the acoustic environment, and how structured sound content in that environment might inform and aid our overall goals, and tasks at hand. Here, I will discuss our group's work (previous and ongoing) on the way that auditory plasticity and learning is driven by the way we interact with the acoustic environment. Our group's work suggests that incidental learning about the structure of our auditory environment influences not only fundamental aspects of perception, but also how we learn to attend and act. Moreover, neuroimaging studies with newly trained participants, as well as expert listeners, have shown the relative utility of a given auditory dimension for a given goal can substantially reshape cortical activity and response preferences across multiple timescales. Work in our group has also unveiled some important counterexamples, for instance where listeners show clear sensitivity to auditory regularities, but fail to use these regularities to improve their task performance, even when task success appears intrinsically rewarding. Conversely, listeners appear to tune responses on task-irrelevant distributional regularities, and in a way that significantly hinders task performance. I will discuss ongoing studies on the neural changes underlying such statistically-driven tuning, as well as some methodological work we have undertaken to facilitate such studies. Finally, and time permitting, I will suggest some ways that our emerging understanding of the dynamism of the auditory auditory might bear on our understanding of functional regionalization and specialization.

ID: 314

Invited talks

Keywords: voice perception, vocal emotion, development, hearing aids, cochlear implants

Auditory development of voice cue and vocal emotion perception in school-aged children with or without hearing devices

Deniz Başkent[1,2]

[1]Department of Otorhinolaryngology/ Head and Neck Surgery, University Medical Center Groningen, University of Groningen, The Netherlands; [2]Research School of Behavioural and Cognitive Neurosciences, University of Groningen, The Netherlands; d.baskent@rug.nl

Hearing aids and cochlear implants have been beneficial in providing auditory information to individuals with hearing loss. The devices do not only amplify the acoustic signals with auditory information to compensate for reduced audibility, they also process and alter them to best compensate for the physiological changes caused by hearing loss. As a result, auditory perception via hearing devices is affected by both compensatory signal processing and hearing loss limitations, and in children, additionally complicated by developmental factors.

Hearing devices have been firstly and mostly developed for delivering speech cues. This focus, combined with neural plasticity, has shown benefit in speech and language development in children with hearing loss and who start using devices at an early age. Human voice is also an important part of speech communication, in conveying speaker information and their affective state, and helping segregating speech from interfering speech. Recently, in a series of studies and using child-friendly game-like test interfaces, we aimed to fully characterize development of voice cue and vocal emotion perception in children of school age (6-18yr), in groups with no hearing loss (baseline development) and with hearing aids or cochlear implants. More specifically, we have measured sensitivity via just-noticeable-differences (JNDs), for the voice cues of voice pitch (F0) and vocal tract length (VTL), and signal detection sensitivity via d' for vocal emotion recognition, for basic emotions with a pseudo-speech carrier.

Our results indicate a strong developmental effect, however, to differing degrees for F0, VTL, and vocal emotions and for each group tested. Even in children without hearing loss, voice cue and vocal emotion perception seem to improve over many years towards adulthood, with differing developmental trajectories. In children with hearing devices, a number of children seem to be able to perceive voice cues at a level expected for their age. However, some children with hearing devices show lower than age-expected scores, and as a result, the group-averaged developmental trajectories may indicate a delay compared to the group with no hearing loss. The voice cue sensitivity scores of children with cochlear implants seem to be better than that of adults with cochlear implants, indicating beneficial effects of early age device use for voice perception. In contrast to voice cue perception, vocal emotion perception seems to be a challenge for most children with hearing devices.

For the few children with lower scores than expected for their age, and for vocal emotion recognition, a remaining question is if the children with hearing loss may catch up to age-expected levels with further device use. On a positive note, our cross-sectional results indicate that hearing devices already work well and seem to facilitate development for some tasks related to voice perception for many device-user children with hearing loss.

ID: 315
Invited talks
Keywords: working memory, memory and learning, relational structure, MEG/EEG

Beyond Chunking: Hierarchical Folding in Auditory Working Memory

Huan Luo
Peking University, China, People's Republic of; huan.luo@pku.edu.cn

Working memory (WM) is constructive in nature, actively integrating incoming information into structural frameworks rather than passively retaining it, thereby overcoming its capacity limitations. A well-established organizational principle in WM is chunking, which groups information into discrete, independent units. In this talk, I will present converging behavioral, neural, and eye-tracking evidence that support a novel structural principle in WM - hierarchical organization. In contrast to chunking that organizes items into isolated slots, hierarchical structure transforms one-dimensional sequences into two-dimensional configurations by aligning substructures and establishing their interrelationships. This hierarchical reorganization not only enhances the efficiency of memory storage and retrieval, but also provides a new framework for understanding WM architecture. Importantly, this principle may extend to other domains of cognition, suggesting its potential generality in learning and memory.

ID: 316

Invited talks

Keywords: auditory cortex, two-photon imaging, learning

The neural basis of flexible auditory learning

Kishore Kuchibhotla

Johns Hopkins University, United States of America; kkuchib1@jhu.edu

I will present recent work demonstrating that the auditory cortex (AC) is necessary for discrimination learning but dispensable at expert levels. Rather than being driven by tonotopic map expansion or tuning shifts, distinct higher-order computations—reward prediction for learning and action suppression for performance—govern behavioral output and are carried by spatially clustered neural ensembles within the AC. These clusters are uncoupled from stimulus-tuning, revealing a higher-order organizational principle of the AC that extends beyond sensory representations. I will then outline future directions in my lab aimed at understanding the neural mechanisms supporting lifelong, multi-task learning, bridging insights from learning theory, the functional architecture of neural circuits, and computational principles of compositionality and generalization.

ID: 317
Invited talks
Keywords: auditory scene analylsis, listening, cortex, ferret

What has neurophysiology taught us about how auditory cortex parses auditory scenes?

Jennifer Bizley
University College London, United Kingdom; j.bizley@ucl.ac.uk

Auditory scene analysis is the process through which the mixture of sounds arriving at ears is organised into a perceptual representation in the brain. While decades of behavioural work have provided insights into the properties of sound that facilitate scene analysis, our understanding of the underlying neural mechanisms remains limited. A key element of auditory scene analysis is the ability to extract temporal structure that unfolds over time. This exists on multiple timescales; for example, we extract short-term statistical regularities, link together individual speech sounds based on pitch continuity or their timbre, and even integrate information from other sensory and motor systems to build stable representations which can ultimately guide behaviour. The auditory cortex is uniquely sited to be play a fundamental role in this process; cells here integrate across frequency and time, are sensitive to context, combine auditory and non-auditory information, and are highly connected to a host of other brain regions. Electrophysiology provides us with the temporal and spatial resolution to understand how neural circuits support auditory scene analysis, but must ideally be combined with behavioural tasks that appropriately engage scene analysis mechanisms. In this talk I will describe some of our attempts to better understand auditory scene analysis across multiple time scales and sound features, considering both spatial and non-spatial representations.

ID: 318
Invited talks
Keywords: auditory cortex, development, subplate, pup, vocalizations

Learning how to hear: The Influence of Early Sensory Experience and Action on Early Cortical Circuits

Patrick Kanold
Johns Hopkins University, United States of America; pkanold@jhmi.edu

Our research aims to understand how experiences and interactions with the world shape the brain. Since hearing is a crucial sense for communication in both humans and animals, we focus on the auditory system. We investigate how sensory representations in the auditory cortex evolve over time and how they are influenced by experience, utilizing in vivo and in vitro imaging and electrophysiological methods. Recent studies from our lab have uncovered insights into the early development and experience-driven plasticity of the auditory cortex.

The brain's ability to process sounds is established during early development, refined through experience during a "critical period," and further shaped by lifelong learning. Our findings reveal that the early developing cortex contains specialized circuits formed by subplate neurons. These circuits provide an early foundation for experience-dependent changes, even before the traditional "critical period."

Our results suggest that both sensory experiences and self-generated actions during early development, before the critical period, can have influences on brain function. Our current research focuses on unraveling the complex nature of these early life experiences and their long-term impact on auditory processing.

ID: 319
Invited talks
Keywords: blindness, visual deprivation, auditory motion, motion, cross-modal plasticity

I can hear what you see: The effects of visual deprivation on auditory processing

Ione Fine[1,2], Woon Ju Park[3]
[1]University of Washington, Seattle, USA; [2]University of Leeds, UK,; [3]Georgia Institute of Technology, Atlanta, USA; ionefine@uw.edu

Almost one-quarter of the brain is normally devoted to processing visual information: reading text, recognizing faces, following the Sunday football match, and much more. In congenitally blind individuals, much of the 'visual' cortex responds strongly to auditory and tactile input rather than to visual stimuli, a phenomenon known as cross-modal plasticity.

Here I will discuss what our laboratory has discovered about the representation of auditory motion perception in early blind individuals, and what this reveals about the plasticity of the human brain.

ID: 320

Invited talks

Keywords: audition, vibrotactile sensation, frequency-following response, MEG, EEG, pitch perception

Periodicity encoding and entrainment across auditory and vibrotactile senses in the human cortex

Emily Coffey

Concordia University, Canada; emily.coffey@concordia.ca

Success perceiving complex auditory stimuli is partly determined by the fidelity with which the auditory system encodes basic properties of sound, such as pitch information. In individuals with hearing loss, poor auditory encoding and processing capabilities increase workload and fatigue, particularly in challenging listening environments. Vibrotactile stimulation offers promising avenues for reinforcing auditory perception, with demonstrable improvements at the behavioural level. However, the neural mechanisms by which improvement is achieved are unclear. This talk will present new work concerning how the auditory and vibrotactile systems process high frequency periodic information. Our recent electroencephalography and intracranial studies reveal both shared encoding mechanisms and distinct neural substrates. We demonstrate that vibrotactile frequency information is encoded in a phase-locked fashion similar to auditory pitch, revealing a distinct vibrotactile frequency-following response. However, cross-modal integration is not observed at this level of encoding, suggesting that sensory integration occurs at a later stage of the sensory processing hierarchy. These neural findings provide context for understanding behavioural benefits of vibrotactile enhancement in hearing-in-noise tasks, while simultaneously raising new questions. For example, "If not at the level of the FFR, where and how can vibrotactile information effectively supplement auditory perception?" The talk will explore how an improved understanding of sensory integration between auditory and vibrotactile modalities may advance both fundamental neuroscience and clinical applications.

ID: 321
Invited talks
Keywords: cross-culture, pitch, rhythm, online, iterated-learning

Mapping cross-cultural internal-representations in music

Kinor Jacoby
Cornell University, Germany; kj338@cornell.edu

How can we overcome the sampling constraint that is entailed by traditional psychology experiments, which recruit participants with access to computer technology from industrialized societies? We apply computational methods and analysis to data obtained in field research with diverse populations around the world in order to investigate whether the similarities we find between participants stem either from biological mechanisms or from comparable exposure. In the first part of the talk I describe how techniques that reduce the reliance on linguistic instruction can be applied to test rhythm perception and production. I will discuss work in which we measured a signature of mental representations of rhythm in 923 participants from 39 participant groups in 15 countries across 5 continents, spanning urban societies and indigenous populations. Integer ratio categories were universally present in rhythm, although their relative importance varied across groups, often reflecting local musical systems. In the second part of the talk I will describe how we use singing to study pitch representations, and specifically results that indicate that octave equivalence is culturally contingent. Finally I will share our recent efforts to extend these lab-based studies to online cohorts. These results demonstrate the exciting possibilities that emerge when we combine computational methods and cross-cultural research to the study of music perception both in the field and online.

ID: 322
Invited talks
Keywords: action, auditory processing

On the impact of action on auditory processing

Erich Schröger
Leipzig University, Germany; schroger@uni-leipzig.de

Brain responses to sound are modulated by endogenous factors like attention and musical expertise. This presentation focuses on how intentional action influences auditory processing, examining three scenarios involving self-generated sounds:

1. Suppression Studies: Self-generated sounds attenuate auditory ERPs (N1, P2) compared to externally generated sounds.

2. Omission Studies: Unexpected omission of action-contingent sounds elicits an omission N1 (oN1), demonstrating that silence can evoke sensory ERPs when sound is predicted. The oN1 is presumably generated in auditory cortex, though it's topography slightly differs from the sound-related N1.

3. Active Oddball Studies: Sounds violating an established physical regularity (deviants) elicit prediction error signals such as the Mismatch Negativity (MMN). When listeners intentionally produce sounds, physically regular (standard) sounds can elicit prediction errors if they mismatch action goals. Vice versa, physically irregular sounds fail to elicit prediction errors if they match action goals.

These findings demonstrating the impact of action on sound processing are compatible with several theories. Motor-to-sensory forward models (efference copy, e.g. von Holst & Mittelstaedt 1950; corollary discharge, e.g., Sperry 1950) state that the predicted consequences of the motor command are sent forward in order to inform the sensory areas about a forthcoming sensory change, which in turn, can be considered by the sensory areas. Also the more general predictive coding theory (e.g. Friston 2021) explains these results quite well. In contrast, however, ideomotor-theory based action selection models (GOALIATH, Hommel 2022); IDEONAMIC, Vogel-Blaschka et al. 2024) can explain the results without recurrence to predictive processes. According to these models, the selection of the action is governed by the action goal. Thus the intended outcome of the action, precedes the selection and execution of the action and no separate prediction mechanism is required for comparing the sensory outcome of the action with the action goal. An extension of the auditory event representation system (AERS) model (Winkler & Schröger 2015) by Korka et al. (2021) is able to integrate predictive and non-predictive comparator models. AERS originally describes the formation of auditory event representations (corresponding to the percepts) as the results of an iterative process of aligning a predictive sensory model (housed in auditory cortex) of what we already know about the acoustic world with the (pre-perceptual) auditory sensory representation driven by the current auditory input. So, in AERS the predictions are derived from a model within the auditory system. According to the extended AERS the predictive auditory model can receive its information not only via bottom-up auditory input, but also via top-down input from non-auditory modules such as the action system. This top-down input must not necessarily be predictive in nature.

ID: 323
Invited talks
Keywords: motor cortex; speech; music; blink

The Role of Motor Simulation and Entrainment in Speech and Music Perception

Yi Du
Institute of Psychology, Chinese Academy of Sciences, China; duyi@psych.ac.cn

This talk explores how the motor system contributes to auditory perception, revealing shared mechanisms between speech and music processing. Using fMRI and TMS, we show that the laryngeal motor cortex actively supports perceptual decisions about lexical tones and consonant voicing, particularly in challenging listening conditions. Extending beyond speech, EEG and eye-tracking experiments demonstrate that spontaneous blink timing—alongside neural oscillations—entrains to musical beat structure, suggesting a novel auditory-motor synchronization mechanism underlying rhythmic prediction. Together, these findings underscore the importance of motor simulation, temporal entrainment, and broader sensorimotor dynamics in shaping auditory perception.

Regular Submissions

ID: 101
Abstract
Keywords: auditory cognition, tonal pitch, MEG, functional connectivity, working memory

Dynamics underlying auditory working memory

Pradeep Dheerendra[1,2], Timothy D Griffiths[2,3]
[1]Glasgow University, United Kingdom; [2]Newcastle University, United Kingdom;
[3]Dept of Imaging Neuroscience, University College London, United Kingdom;
pradeep.dheerendra@gmail.com

We aim to understand the dynamics underlying auditory working memory for maintaining 'simple' tones. We recorded magnetoencephalography (MEG) in 17 subjects while they maintained one of the two presented tones (or ignore both in the control condition). After 12s, subjects compared the pitch of a test tone with the maintained tone.

Analysis of evoked responses showed persistent activity throughout maintenance compared to the pre-stimulus silent baseline but only at the start of maintenance when compared to the control condition. The evoked response during maintenance was source localised against baseline to bilateral auditory cortex. Analysis of induced responses showed suppressed alpha in the left auditory cortex, enhanced theta in medial prefrontal cortex, and enhanced beta in cerebellum.

In a second experiment, 19 new subjects were presented with a tone and a Gabor patch and a retro-cue indicating whether to maintain auditory or visual information for 12s. Analysis of the induced responses in auditory condition yielded results similar to those observed in the first experiment.

Connectivity analysis showed that the theta activity in medial prefrontal was phase-locked to activity in the left hippocampus and left auditory cortex. The beta activity in cerebellum was phase-locked to left Inferior Frontal Gyrus (IFG) activity and correlated to subject's task accuracy.

Using MVPA, a LDA classifier was trained to decode the contents of AWM (discriminate b/w low vs high pitched tone) using beta band Phase Locking Value with right cerebellum as its features. A channel searchlight analysis showed that decoder performance at right Anterior Cingulate Gyrus (56.17% acc.) was above chance. Further, the decoder performance at Right STG & MTG was correlated (rho=0.725, p<0.01) to subject's task accuracy, showing a correspondence between encoding distance and behavioural performance.

Our data clearly shows a network of brain areas involving pre-frontal and hippocampus, IFG and cerebellum for maintaining sounds in the auditory cortex, consistent with previous fMRI [1] and ECoG experiments [2].

References:

[1] Kumar, S., Joseph, S., Gander, P.E., Barascud, N., Halpern, A.R. and Griffiths, T.D., 2016. A brain system for auditory working memory. Journal of Neuroscience, 36(16), pp.4492-4505.

[2] Kumar, S., Gander, P.E., Berger, J.I., Billig, A.J., Nourski, K.V., Oya, H., Kawasaki, H., Howard III, M.A. and Griffiths, T.D., 2021. Oscillatory correlates of auditory working memory examined with human electrocorticography. Neuropsychologia, 150, p.107691.

ID: 102

Abstract

Keywords: voice; acoustic patterns; voice area; acoustic patterns; fMRI

Voice-sensitivity in auditory cortex is based on acoustic pattern decoding

Sascha Frühholz[1,2], Matthias Staib[2]

[1]University of Oslo, Norway; [2]University of Zurich; s.fruehholz@gmail.com

A subregion of the primate auditory cortex (AC) was proposed to selectively process voices and voice sounds, but the detailed functional processing mechanisms and the proposed selectivity of this voice area (VA) remain rather elusive. For a better functional description of the VA, we investigated its neural responses to vocal sounds on one side and non-vocal acoustic patterns and features on the other hand. In two experiments, we critically compared neural AC responses to original voice sounds with neural responses to simple textural sound patterns (TSP) that share basic acoustic features with natural sounds but that are perceptually very distant from voice sounds. We found, first, that listening to simple TSP elicited activity in large regions of the VA, driven by the variance in acoustic features of TSP that typically also differ between voice and non-voice sounds. Second, we reconstructed the VA activity that is usually found for voice processing from activation patterns elicited by simple TSP. The reconstruction model specifically shows that the VA equally parameterizes both voice and non-voice sounds by their acoustic and perceptual features. Third, a machine-learning approach confirms that the simple TSP contains almost all the information necessary to mechanistically explain VA activity for voice processing. And fourth, by eliminating activity in the VA that is largely driven by basic sound features rather than higher-level voice precepts, it turns out that only a minor fraction of the original VA could potentially be determined as “voice-selective”. Taken together, rather than being a selective brain node reserved only for voice processing, the largely extended human VA seems sensitive to a much broader range of sounds and uses mechanisms to non-selectively evaluate the acoustic and perceptual patterns similarly for both environmental sounds and for voice sounds.

ID: 103
Abstract
Keywords: EEG, auditory, perception, N100, P200

When do we hear?

Maysan Bader[1], Leon Deuoell[1,2], Israel Nelken[1,3]
[1]Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Israel, Israel; [2]Department of Psychology, The Hebrew University of Jerusalem, Jerusalem, Israel; [3]Department of Neurobiology, the Silberman Institute of Life sciences, Hebrew University of Jerusalem, Israel;
maysan.bader@mail.huji.ac.il

When Do We Hear?

Maysan Bader

The perceptual center (P-center) of a sound is defined as the moment in time at which a sound is perceived to have occurred. P-centers can be estimated experimentally by Alignment tasks (aligning target sound with a click) or Tapping tasks (tapping to an isochronous beat of target sound). Previous studies have shown that the p-center is affected by sound structure, proficiency in music, the requirements of the task, and many other factors. Here, we relate the P-center to the neural mechanisms subserving auditory perception. We estimated P-centers for a set of previously tested sounds using a temporal order judgment paradigm, in which participants heard two short periodic sound sequences concurrently: one consists of the target sound, and the other consists of clicks. The two sequences were shifted relative to each other (Stimulus Onset Asynchrony, SOA), and the listeners were asked to report whether the clicks preceded or followed the target sounds. The P-center was defined as the SOA corresponding to a chance decision.

EEG was recorded simultaneously to examine the relationship between the behavioral P-center and neural auditory processing. A Linear Mixed Effect model was used to estimate the behavioral P-center as a function of the N100 and P200 peak latencies of the ERP evoked by the target sound. The latencies of the ERP components explained a small yet significant within-subject variation. However, most of the variability of the P-center was due to cross-subject differences. Remarkably, there were significant correlations between subject-specific ERP measures and the P-center, including a negative correlation with N1 latency and a highly significant negative correlation with the alpha peak frequency. We, therefore, suggest that earlier auditory perception could be related to faster brain processes.

ID: 104
Abstract
Keywords: auditory perception, representational map, decision-making

Representational geometry of pulsed sound stimuli in the auditory cortex explains perceptual decision-making in mice

Johannes Seiler[1], Giuseppe Cazzetta[1], Takahiro Noda[1,2], Simon Rumpel[1]
[1]Institute of Physiology, Focus Program Translational Neurosciences, University Medical Center of the Johannes Gutenberg University Mainz, Germany; [2]Graduate School of Medical Sciences, Kyushu University, Higashi-ku, Fukuoka, Japan; johseile@uni-mainz.de

Our brain is thought to encode sensory information, such as incoming sound stimuli, in form of representational maps, reflecting relations between different sensory inputs. While a representational map's structure effectively captures different features of sensory stimuli, it remains unclear how this structure informs perceptual discrimination. Here, we addressed this issue in the mouse auditory system: We first tested perceptual discrimination of a broad set of pulsed sound stimuli in a categorical go/no-go task that requires the auditory cortex for successful discrimination. We observed that mice showed consistent biases in their choices for stimuli that are the same with respect to the reinforced feature (number of pulses), however, differ in other temporal aspects (e.g. clusteredness or irregularity). Next, we applied large-scale two-photon calcium imaging of thousands of neurons to capture the global structure of stimulus-evoked activity in the auditory cortex. Mapping the population activity to a parametric framework of an interpretable set of stimulus features, we identify those features prominently represented in the representational map of the auditory cortex. We find that systematic choice biases can be explained by a representational entanglement of reinforced and non-reinforced stimulus features. Together, our findings highlight the representational architecture in the auditory cortex as a relevant underpinning of perceptual decision-making.

ID: 105

Abstract

Keywords: Auditory Plasticity, Auditory Training, Speech-in-noise, fNIRS, MTG

Tracking Auditory Plasticity: fNIRS Insights into Rapid Neural Reorganization Following Brief Auditory Training

Yael Zaltz, Stav Bracha

Tel Aviv University, Israel; yaelzalt@tauex.tau.ac.il

Speech perception in noise (SIN) is critical for communication, yet its neural plasticity mechanisms remain incompletely understood. This study used functional near-infrared spectroscopy (fNIRS) to examine cortical reorganization following a single-session SIN training and assess the reliability of fNIRS as a tool for tracking auditory learning. Thirty normal-hearing adults (19–36 years) underwent three sessions spaced 1 to 3 days apart. The first session included a hearing test, fNIRS data collection, and a 45-minute SIN training, with six adaptive assessments of speech reception thresholds in noise (SRTn). In the second session, two SRTn assessments evaluated the behavioral impact of training, while fNIRS data assessed its physiological effects. The third session exclusively focused on fNIRS data collection to evaluate test-retest reproducibility. All stimuli for testing and training underwent spectral degradation via an eight-channel noise vocoder. fNIRS data were collected using a pseudorandomized block design, alternating among three conditions: speech stimulation, speech-shaped noise stimulation, and speech-in-noise stimulation at a signal-to-noise ratio of +5 dB. Activation was assessed via 41 fNIRS channels, positioned over the temporal, prefrontal, and frontal cortical regions. Hemodynamic responses (HBO and HBR beta values) were extracted using Homer 3 and analyzed with a general linear model (GLM). These values were compared across conditions and sessions and correlated with behavioral improvements in SRTn. Results demonstrated significant improvements in SIN perception following training, particularly in participants with poorer initial performance or lower pre-training activation in the left middle temporal gyrus (MTG) in response to the stimuli. A significant post-training reduction in left MTG activation for trained stimuli accompanied this behavioral enhancement. Given the MTG's role in auditory analysis and speech decoding, this decrease may reflect the automation of speech processing and enhanced suppression of irrelevant noise, suggesting increased neural processing efficiency. Despite robust group-level effects, individual-level reproducibility of fNIRS activation patterns was low, posing challenges for clinical translation. These findings underscore the potential of fNIRS in mapping auditory plasticity but highlight the need for methodological refinements to enhance test-retest reliability. Understanding cortical adaptation to brief auditory training may inform targeted interventions to improve speech perception in challenging listening environments.

ID: 106

Abstract

Keywords: EEG, Envelope Tracking, Hearing Aids, Cochlear Implants

Divergent Cortical Speech Tracking at the Cocktail Party: Preserved in Hearing Aid Users, Impaired in Cochlear Implant Users

Constantin Jehn[1], Niki Katerina Vavatzanidis[2], Anja Hahne[2], Tobias Reichenbach[2]

[1]Friedrich-Alexander Universität Erlangen, Germany; [2]Ear Research Center Dresden (ERCD), University Hospital Carl Gustav Carus; constantin.jehn@fau.de

While hearing aids (HAs) or cochlear implants (CIs) can often restore speech perception in quiet, background noise and competing speech remains a challenge, requiring high listening effort or preventing conversation for hearing impaired individuals altogether. Neurofeedback, using EEG to identify the attended talker, could optimize HA and CI processing to enhance speech comprehension in noise [1]. Developing such devices requires a deeper understanding of neural speech tracking in hearing impaired individuals. This study compares cortical tracking of the speech envelope among HA users, CI users, and an age-matched control group in competing speaker situations.

To this end, we collected EEG data from 25 bilateral CI users, 29 bilateral HA users, and 29 age-matched controls in a selective attention paradigm. Our study comprised eight trials with one audiobook presented (single speaker) and 12 trials with two audiobooks presented from different directions (competing speaker). Each trial lasted two minutes and the audiobooks were presented through loudspeakers. The participants were instructed visually on which story to focus. To quantify the participants' understanding of the presented audiobooks, comprehension questions were posed after each trial. Additionally, we assessed the subjective listening effort under the different listening conditions.

We used a linear forward model to study the neural response to the speech envelopes. A linear backward model served to examine the differences in reconstruction scores of the attended speech envelope and the decoding accuracy. One-way ANOVA was used to test for group differences. If applicable, post-hoc comparisons were carried out using unpaired t-tests. All p-values were FDR-corrected.

Behaviorally, we found significant group differences only in the competing speaker paradigm. Subjective listening effort ratings increased successively from controls to HA users to CI users (all $p<0.01$). On the other hand, comprehension scores were significantly reduced in the CI group only ($t=6.91$, $p<0.001$), as HA users and controls both performed at ceiling.

The decoding accuracy of the attended speaker was significantly reduced for the CI group ($t=6.31$, $p<0.001$) . This could be explained by higher reconstruction scores of the distractor envelope ($t=4.84$, $p<0.001$) that we observed for the CI group. Also, the HA group showed increased reconstruction scores of the distractor speaker compared to the control group ($t=2.01$, $p=0.048$). No significant group differences in reconstruction scores of the attended speech envelope emerged.

This study demonstrates that auditory attention can be reliably decoded from the speech envelope, also for participants with hearing impairment who wear auditory prostheses. Our study further provides evidence that speech perception difficulties in cocktail party environments among hearing-impaired listeners are linked to impaired cortical suppression of the distractor speaker.

ID: 107
Abstract
Keywords: Eye movement-related eardrum oscillations; saccades; sensory-motor integration; eye movement; middle ear

Relation of eye movement-related eardrum oscillations (EMREOs) amplitude and time course with saccade direction and tympanometric measurements

Nancy Sotero Silva[1], Christoph Kayser[1], Felix Bröhl[2,3]
[1]Universität Bielefeld, Germany; [2]Max Planck Institute for Human Development, Berlin, Germany; [3]Max Planck Dahlem Campus of Cognition, Berlin, Germany; n.sotero@uni-bielefeld.de

Eye movement-related eardrum oscillations (EMREOs) are rhythmic signals recorded by in-ear microphones and that are systematically aligned to saccadic eye movements. Those signals last for up to 100 ms and reflect movements of the tympanic membrane that scale with the magnitude and direction of saccades, showing opposing phases for saccades towards locations ipsi-and contralateral to the ear in focus. Yet many questions regarding this phenomenon remain unresolved. Based on systematic and statistical analysis of bilateral in-ear recordings in human participants we here explore several properties of these EMREOs in order to improve our understanding of this signal's origin and functional significance, focusing on four specific questions: (1) comparison of potential differences in the EMREO amplitude time course between saccade directions; (2) comparison of EMREO time course between left and right ears; (3) comparison of EMREOs induced by visually and auditory guided saccades; (4) and tests for a relation between the EMREO amplitude and middle ear properties (assessed by tympanometry, measurement of acoustic reflexes and audiometric testing). For that purpose, eye tracking and in-ear data from 24 subjects were recorded during horizontal saccades guided for visual and auditory targets (from -18 to 18 degrees azimuth). We found that the EMREO time course is comparable between the left and right ears, and between visual and auditory guided saccades. On the other hand, ipsi- and contralateral saccades amplitude time course present a significant difference. Our results also support that the EMREO amplitude is negatively related to the compliance of the tympanic membrane and to the equivalent ear canal volume. Together, these results suggest that EMREOs reflect motor-related top-down signals, and fuel the speculation that EMREOs may be generated by the middle ear muscles in a differential operation similar to the execution of ipsi-and contralateral saccades.

ID: 108

Abstract

Keywords: Auditory Discrimination Learning, Neuronal Discriminability, Feature Abstraction, In Vivo Extracellular Electrophysiology, Primary Auditory Cortex

Curriculum Learning: sequential acquisition of task complexity enhances neuronal discriminability

Maria Shujah, Tania Rinaldi Barkat

University of Basel, Switzerland; tania.barkat@unibas.ch

Throughout history, learning strategies have been optimized to enhance efficiency. One particularly effective strategy is Curriculum Learning (CL), where tasks of increasing complexity are acquired sequentially. This strategy, commonly used in educational and machine learning contexts, accelerates learning. While this training strategy has been widely studied on a behavioral level, the neural underpinnings of CL have remain largely unexplored. Our study addresses this gap by systematically investigating the behavioral and neural mechanisms underlying CL using a novel auditory discrimination paradigm in mice.

To characterize CL behavior, we designed a multi-staged appetitive Go/No-Go learning paradigm where a hard auditory discrimination task was preceded by an easier one. Both tasks required the mice to discriminate between a pulse train with a single tone (No-Go) and a pulse train with alternating tones (Go). The key difference between the tasks was the nature of the tones: the easy task used pure frequency tones (PT), while the hard task used harmonic complex tones (HCT). At the behavior level, we found that PT discrimination was indeed easier than HCT discrimination. To study the effect of CL, we created three cohorts: one cohort underwent progressive training from the easy task to the hard task (PT-HCT group). In contrast, the two other cohorts were subjected to the easy task only (PT group) or to the hard task from the outset (HCT group). After training, we performed in vivo extracellular electrophysiology in the auditory cortex of the mice while they passively listened to the stimuli.

Our results demonstrate that sequentially structuring auditory learning tasks with increasing complexity significantly enhances both the speed and accuracy of learning. In vivo electrophysiology revealed that neuronal discriminability of sounds improves with training. Interestingly, training on the easy task alone enhances discriminability for the hard task, at both single-cell and population levels, which indicates a transfer of learning. Furthermore, within the PT-HCT group, the differences between Go and No-Go responses for tones used both in the easy and hard tasks became highly similar, suggesting a convergence on common discriminative features.

Our study uncovers critical neural mechanisms that facilitate learning through progressively complex tasks. We discover that neurons in the primary auditory cortex gradually abstract task parameters, becoming more selective and focusing on relevant features as training progresses in the direction of increasing complexity. This process expands on schema formation, where the brain organizes and interprets information efficiently, reducing computational load and improving learning performance.

ID: 110
Abstract
Keywords: Otoacoustic Emissions, Auditory Attention Decoding

Attentional Modulation of Otoacoustic Emissions Evoked by Speech-Derived Stimuli: Examining the Role of Resolved and Unresolved Harmonics

Janna Steinebach, Tobias Reichenbach
Friedrich-Alexander-University Erlangen-Nürnberg, Germany;
janna.steinebach@fau.de

The human auditory system excels at selectively focusing on a sound of interest while suppressing background noise, a process essential for understanding speech in complex acoustic environments. This may be aided by efferent feedback from the auditory cortex to the inner ear, or cochlea. The efferent feedback may indeed be able to regulate the amount of mechanical amplification in the cochlea in specific frequency bands to aid selection of a target voice.

A key phenomenon linked to the cochlea's active process are otoacoustic emissions (OAEs), faint sound waves generated by the basilar membrane's nonlinear response to sound. The so-generated mechanical disturbances travel retrogradely through the cochlea, and emerge as measurable sound from the ear canal.

OAEs are classified based on their eliciting stimuli. For example, upon stimulation with two frequencies, so-called distortion product otoacoustic emissions (DPOAEs) arise at an additional frequency that results from linear combinations of the originals. To date, DPOAEs have almost exclusively been measured in response to pure tones.

In our study we employ more complex stimuli based on natural running speech to measure speech-OAEs. Extending our earlier work on this topic, we consider stimuli that are based on multiple harmonic overtone pairs of the voiced parts of speech [1]. We thereby include both resolved and unresolved harmonics in order to assess cochlear activity and putative effects of attention. As resolved harmonics produce distinctive peaks on the basilar membrane whereas unresolved harmonics peak in overlapping areas, we suspect that effects of attentional modulation are more pronounced for resolved than for unresolved harmonics. Furthermore, we designed the stimuli such that simultaneous measurements of speech-OAEs corresponding to a male and to a female voice is possible.

We then measured speech-OAEs during selective auditory attention. Subjects heard two competing voices, a male and a female one, in one ear, while speech-OAEs were measured from the contralateral ear. Participants were instructed to focus on one voice at a time, with the focus of attention shifting regularly between the two speakers.

In total, 40 participants aged 18 to 31 took part in the study. First results indicate that the attentional focus indeed modulates the morphology of DPOAEs. Further analysis will give insight into the extent of differences between resolved and unresolved harmonics and their degree of modulation due to attentional effects.

Our results show that speech-DPOAEs can be measured reliably both for resolved and unresolved harmonics. They constitute a promising tool for probing the role of efferent feedback to the inner ear in the cognitive aspects of speech processing, such as selective auditory attention.

[1] Saiz-Alía, M., Miller, P., & Reichenbach, T. (2021). Otoacoustic emissions evoked by the time-varying harmonic structure of speech. Eneuro, 8(2).

ID: 111

Abstract

Keywords: Ultra-high-field fMRI, Magnetoencephalography (MEG), Predictive processing, Repetition suppression

Dissociating Contributions of Expectation and Repetition Suppression Using High-Field fMRI and MEG.

Jorie van Haren[1], Sonja Kotz[1], Floris de Lange[2], Federico De Martino[1]

[1]Maastricht University, Netherlands; [2]Donders Institute, Radboud University Nijmegen, Netherlands; jjg.vanharen@maastrichtuniversity.nl

To make sense of the soundscape of our surroundings, listeners continuously use contextual information to form predictions about what is likely to occur next while suppressing repeated sensations. Although both prediction (Tang et al., 2021) and repetition suppression (Todorovic & de Lange, 2012) aid neural sound processing, differentiating between the two remains challenging. Surprising events often coincide with changes in low-level attributes, triggering a release from adaptation. Despite their co-occurrence, expectations may modulate neural responses beyond what can be explained by repetition suppression alone.

Here, we investigate how past experiences, present stimuli, and inferences towards the future shape auditory perception by presenting probabilistic tone sequences and employing ultra-high-field (7T) fMRI to examine layer-specific effects of tuning, expectations, and repetition suppression. Complementarily, we used magnetoencephalography (MEG) to discern the temporal dynamics of the repetition suppression–prediction interplay. Our content-specific, tuning-dependent modeling approach included three components: 1) low-level tuned activation, reflecting expected voxel activation patterns based on population receptive field dynamics; 2) expectations, quantified using the DREX Bayesian observer model (Skerritt-Davis & Elhilali, 2021); 3) and repetition suppression, captured by a double-exponential decay model accounting for both rapid adaptation to near-repetitions and lingering effects across multiple stimuli (Fritsche et al., 2021). Time-resolved regressions were used to derive temporal response functions for each component in the MEG data.

In early auditory cortex, including Heschl's gyrus and planum temporale, tuning, repetition suppression, and expectations overlapped, whereas later regions such as planum polare and superior temporal gyrus showed distinct contributions by each component. Expectations explained variance beyond tuning and repetition suppression, with the strongest effects in deep cortical layers and minimal effects in superficial layers. Temporally, MEG revealed a rapid-onset, sustained repetition suppression response, followed by priors peaking and dipping in time. Finally, surprisal responses resembling Mismatch Negativity signaled expectation violations. These laminar and temporal patterns suggest distinct computational roles: while repetition suppression facilitates rapid adaptation to stable input, priors and surprisal highlight deviations, driving model updates. Together, our findings illustrate how past experiences (repetition suppression), present input (tuning), and future predictions (priors, surprisal) interact to balance stability and sensitivity to change, shaping auditory perception across cortical hierarchies.

ID: 112

Abstract

Keywords: binaural, spectro-temporal, natural sounds, encoding model, ferret

Topography of binaural spectro-temporal tuning in primary auditory cortex

Alexis Zou[1], Jereme C. Wingert[2], Stephen V. David[1]

[1]Department of Otolaryngology, Oregon Health & Science University; [2]Behavioral and Systems Neuroscience Program, Oregon Health & Science University; zoua@ohsu.edu

It is generally agreed that single-unit activity in the auditory cortex (AC) encodes information about both spectro-temporal and binaural sound features. However, previous research has typically focused on just one of these areas and few have explored the binaural representation of spectro-temporal features. It has been established that some spectral and binaural tuning properties, like best frequency and contralateral-ipsilateral balance, are topographic; that is, anatomically adjacent neurons tend to share these tuning properties. This study aimed to establish a deeper understanding of the topography of joint spectrotemporal and binaural tuning, specifically asking how binaural sound features are processed in the auditory cortex and if certain binaural parameters are associated with distinct cortical columns. Linear electrode arrays were used to record the activity of 3991 neurons from primary and secondary AC (67 recording sites in 6 awake, passively listening animals) during the presentation of two simultaneous natural sound streams, presented +/-30 degrees azimuth from the midline. For each neuron, encoding model analysis was used to fit a binaural STRF, an extension of the traditional, single-channel STRF, to account for responses to two spatially distinct stimuli. We compared three model architectures of binaural encoding: the monaural model collapsed two sounds into a single input spectrogram, the allocentric model treated the two sounds as distinct, parallel inputs, and the HRTF model used the head-related transfer function to compute the sound reaching each ear and used those spectrograms as inputs. On average, the HRTF model was more accurate than the allocentric and monaural models, consistent with a predominantly egocentric representation of sound in A1. Both spectrotemporal and binaural tuning properties were often more similar between neurons in the same cortical column (i.e., the same recording site) than in different cortical columns. In many neurons, we observed different spectral tuning between ipsilateral and contralateral ears. The presence or absence of binaural tuning differences was also correlated within cortical columns. In addition to encoding sound location, the joint encoding of spectro-temporal and binaural sound features may support grouping and stream segregation in complex auditory scenes.

ID: 114
Abstract
Keywords: Hearing loss, Brain stimulation, Neuroprosthetics

Cortical implants enhance hearing restoration and mimic natural auditory processing more closely than cochlear implants

James Alexander Taylor[1], Natasa Jovanovic[1], Charlotte Navntoft[1], Amelie Albon[2], Shahab Rezaei-Mazinani[2], Brice Batthelier[3], Tania Rinaldi Barkat[1]
[1]University of Basel, Switzerland; [2]Mines Saint-Etienne, France; [3]Institut Pasteur, France; jamesalexander.taylor@unibas.ch

Hearing loss affects half a billion people worldwide, significantly impairing communication and quality of life. While hearing aids provide adequate support for mild cases, severe hearing loss often requires cochlear implants. Despite their success, cochlear implants suffer from poor spectral accuracy, low acoustic resolution, and ineffectiveness in patients with auditory nerve damage. This study investigates whether cortical implants offer a more effective alternative for auditory restoration.

We implanted surface electrode arrays with 32 contacts (175 µm spacing) onto the auditory cortex of mice and compared them with cochlear implants featuring 4 electrodes (500 µm spacing). To ensure a direct comparison, we analyzed responses from four cortical electrodes, four cochlear implant electrodes, and four frequency-matched acoustic stimuli in normal-hearing mice. Mice were tested on single- and multi-frequency categorization tasks, amplitude modulation categorization, and amplitude modulation in noise categorization. Across all tasks, cortical implants led to superior auditory performance compared to cochlear implants, although learning took longer and followed distinct learning curves. Notably, mice trained with natural sounds could perform the tasks with cortical stimulation without additional training and vice versa, suggesting that cortical stimulation evokes percepts that closely resemble natural hearing.

To investigate the neural basis of these differences, we recorded in vivo electrophysiological responses in the auditory cortex to cortical stimulation, cochlear stimulation, and natural sounds. Neuronal responses scaled with stimulus strength across all conditions. Dimensionality reduction analyses (PCA, t-SNE) revealed that cortical stimulation responses resembled natural sound-evoked responses more closely than cochlear stimulation responses. Linear decoder analyses trained on natural sounds and tested on cortical and cochlear stimulation showed higher classification accuracy for cortical stimulation. Additionally, three-way decoder analyses confirmed that cortical stimulation responses were more similar to natural sounds than to cochlear stimulation, aligning with behavioral findings.

These results suggest that cortical stimulation not only provides better auditory performance but also more faithfully mimics natural auditory processing. The ability of mice to generalize between natural sound and cortical stimulation further supports this claim. While cortical implants show promise as an alternative for hearing restoration, future studies should further investigate learning dynamics, perceptual fidelity, and translational feasibility for clinical application.

ID: 115

Abstract

Keywords: ventral tegmental area, music preference, dopamine, music sounds, fiber photometry

Early music exposure affects neural activity in the ventral tegmental area (VTA)

Samira Souffi, Israel Nelken

Hebrew University of Jerusalem, Israel; samira.souffi@mail.huji.ac.il

Dopaminergic release from neurons of the ventral tegmental area (VTA) has been widely studied for its role in reward prediction error encoding, behavioral reinforcement, motivational salience and learning processes. While dopamine release has been implicated in determining sound preferences, this role is less well characterized. We exposed male and female mice in their homecage to human music (1st movement of Beethoven's 9th symphony) or silence from P7 to P40, covering both early and late auditory critical periods. At early adulthood, we performed a free-choice behavioral test in which mice could choose to dwell in a music zone or in a silence zone of the test box. Following the test, we performed fiber photometry in the VTA of these mice, while they moved freely and heard simple (broadband noise and pure tones) and complex (exposed and unexposed music excerpts) sounds. On average, we found that exposed mice (music- or silence-exposed) spent a longer time in the music zone than in the silence zone compared to naive mice, with noticeable differences between males and females. All sounds were associated with both excitatory and inhibitory transients. The music-exposed mice showed a large decrease in VTA activity to all sounds compared to naive mice; silence-exposed mice showed a decrease in activity (males) or no significant change (females). Finally, in naive males, we revealed a positive relationship between VTA activity and the time spent in the music area. Further experiments are currently conducted on naive TH-cre mice using GCaMP virus allowing chronic recordings of the dopaminergic neurons only.

Overall, these data indicate that early sound experience had a strong effect on music preference behavior as well as on VTA activity. Whether and how the changes in activity in VTA affect music preferences in the exposed mice needs further work.

ID: 116
Abstract
Keywords: sensorimotor, audition, auditory decoding, electrophysiology

Action and Perception in the context of a sensorimotor task in the ferret

Felicie Levinton[1], Yves Boubenec[1], Shihab Shamma[2,1], Flavien Feral[1]
[1]Ecole Normale Supérieure, PSL University, Paris, France; [2]Institute for Systems Research, University of Maryland, College Park, MD 20742, USA;
levintonf@gmail.com

The ability to predict and differentiate self-generated (reafferent) sounds from external auditory stimuli is crucial for detecting important environmental cues and is fundamental for mastering complex acoustic behaviors such as speech production in humans or vocalizations in animals. Central to this capability is the intricate coordination of motor and sensory neural processes, allowing precise control and anticipation of acoustic consequences of movements.

Our study investigates the neural circuits underlying sensorimotor interactions involved in learning to anticipate self-produced sounds and explores how auditory cortical responses adapt through implicit sensorimotor learning. To achieve this, we developed a behavioral setup in which the azimuthal position of the ferret head is coupled to the frequency of acoustic feedback in a closed-loop configuration. Two conditions were compared: a "tracking" (closed-loop) condition where the acoustic frequency was synchronized with the ferret’s head position, and a "playback" (open-loop) condition lacking this movement-sound correlation.

Using simultaneous electrophysiological recordings from both primary and secondary auditory cortices during task performance, we observed significant modulation in cortical responses to reafferent sounds. Specifically, neural activity in both primary and secondary auditory cortices exhibited a substantial reduction during the tracking condition compared to playback, indicating selective suppression of anticipated sensory consequences of self-generated movements. Interestingly, this suppression was more pronounced in the primary auditory cortex compared to the secondary auditory cortex. Additionally, we identified progressive modifications in auditory cortical responses across multiple training sessions, indicative of implicit learning of the motor-sensory mapping.

These findings illustrate how motor-related signals can dynamically reshape auditory cortical processing, facilitating the selective attenuation of predictable acoustic feedback. Our results provide compelling evidence for the crucial role of sensorimotor interactions in auditory cortical plasticity and highlight the underlying neural mechanisms central to sensorimotor learning and control.

ID: 117

Abstract

Keywords: Sound localization, spatial attention, rapid plasticity, neural decoding, auditory cortex

Task engagement modulates spatial coding in the ferret auditory cortex

Waradon Phokhinanan[1], Jonatan Nordmark[2], Shihab Shamma[3], Yves Boubenec[1]

[1]Laboratoire des systèmes perceptifs, Département d'études cognitives, École normale supérieure, PSL University, Paris, France; [2]Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Solna, Sweden; [3]Neural Systems Laboratory, University Of Maryland, Maryland, USA; waradonai@gmail.com

Spatial location is a fundamental feature of sound perception, yet the neural mechanisms underlying selective spatial attention in the auditory cortex remain unclear. While frequency-driven attention has been shown to induce rapid cortical plasticity in spectro-temporal receptive fields during tone detection tasks in ferrets, studies in the cat auditory cortex have primarily reported general sharpening of spatial tuning during task engagement. Here, we investigate whether spatial attention can selectively modulate auditory source encoding at the attended location through behavior-dependent rapid plasticity.

To address this question, we conducted Go/No-Go experiments on two female ferrets using a head-fixed sound position change-detection task in an anechoic chamber equipped with 12 speakers arranged along the horizontal plane at 15-degree intervals, covering a 180-degree frontal field. In each trial, reference stimuli were presented 2–4 times from a randomly selected speaker, with all reference sounds originating from the same location. In some trials, a target stimulus was introduced at a different location. The ferrets were trained to detect this change and respond by licking at the target position to receive a reward, while incorrect responses to reference stimuli resulted in timeouts. Both reference and target stimuli consisted of broadband noise (125–4000 Hz) with randomly varying center frequencies, ensuring that no two sounds in a trial were identical. Neural activity was recorded using a 64-channel Cambridge Silicon Probe (M1), capturing single-unit spiking responses across ~7000 trials from 900 neurons in both behaving and passively listening conditions.

Behavioral results indicate that both ferrets exhibited selective attention to specific sound locations, as reflected by their highest accuracy. This effect was independent of the spatial distance between reference and target positions, as well as trial duration, which varied due to the random number of reference sound repetitions. Using single-neuron activity, we successfully decoded cue and target positions. Notably, we found that spatial position encoding was enhanced during task engagement. Individual neurons exhibited improved spatial coding for specific positions, suggesting a localized and heterogeneous effect of selective spatial attention. Future analyses will incorporate intermittent probe clicks (~50 ms) presented from the 12 speakers across trials to further characterize attention-driven changes in spatial tuning.

ID: 118
Abstract
Keywords: category learning, feedback intervention, human, real-time imaging

Towards real-time policy inference in multidimensional auditory category learning

André Brechmann, Jörg Stadler, Nicole Angenstein, Marcel Lommerzheim, Susann Wolff
Leibniz Institute for Neurobiology, Germany; brechmann@lin-magdeburg.de

Since decision making is the result of internal brain processes, observing the actions of a learner in complex problem solving is not sufficient to identify the underlying behavioral policy. Thus, measuring brain activity and psychophysiological data may help to disambiguate the interpretation of externally observable behavior. Such combined information could then be used for providing supportive feedback intervention while preventing premature interruption of the ongoing learning process. This in turn allows to study the impact of such intervention on subsequent learning and brain activity in order to unravel the neural correlates of auditory learning.

Following such an approach, we developed an experimental setup that simultaneously measure whole brain fMRI, EEG, psychophysiological and detailed behavioral data while participants perform a multidimensional auditory category learning task. A large set of sounds with five different features of two possible values each are presented in 180 trials: duration (short/long), pitch direction (rising/falling), loudness (low/high), frequency range (low/high), and speed of pitch change (slow/fast). Participants were not informed that the target category is defined by the features duration and direction (e.g. long, rising) but had to learn it by trial and error. For each trial, they had to indicate via button press whether they thought the sound belonged to the target category (right index finger) or not (right middle finger) and received verbal feedback about the correctness of the response. A study on feedback-related processes using this paradigm (Wolff & Brechmann 2023) and a new set of data on 24 subjects was used to identify brain regions that serve as volumes of interest for real-time fMRI analysis with Turbo BrainVoyager on 34 additional subjects.

Offline analysis of pilot data identified brain regions that develop differential learning-dependent activation for target vs. non-target trials, i.e. auditory association cortex, anterior cingulate, medial/dorsolateral prefrontal cortex, anterior insula, caudate, dorsomedial thalamus, subthalamic nucleus, dopaminergic midbrain. More detailed analysis of learning performance revealed a subset of participants who persistently stick to a one-feature policy while ignoring the second target feature. Activation of the identified brain network was found to reflect the different policies which suggests that patterns of fMRI activation can help to identify the participants' policy. Real-time analysis of the dynamics of brain activity in single subjects and introspective reports of applied strategies, however, reveal a much wider range of variance and dynamic of behavioral policies, including visual or verbal strategies with concomitant activation of visual resp. inferior frontal cortex. Our contribution will discuss this issue of increased dynamics and inter-individual variability of brain activation when studying more complex auditory category learning paradigms.

ID: 119

Abstract

Keywords: Predictive Auditory Processing, Predictive Coding, Audio-Visual Paradigm, high-resolution fMRI

The Laminar Patterns of Predicting Content in Learned Audio-Visual Associations

Mahdi Enan[1], Agustin Lage Castellanos[1], Ryszard Auksztulewicz[1,2], Federico De Martino[1]

[1]Maastricht University, Maastricht, The Netherlands; [2]Free University Berlin, Berlin, Germany; m.enan@maastrichtuniversity.nl

To make sense of the external world, our brains use contextual information to predict upcoming stimuli. Previous studies have shown that content ("what") and temporal ("when") predictions are supported by different mechanisms (Auksztulewicz et al., 2018). In particular, "what" predictions are linked to stimulus-specific gain modulation. Here, we aimed to understand how these mechanisms are grounded within the laminar cortical architecture using laminar fMRI. Within the predictive coding framework, we hypothesized that the laminar responses for prediction errors increase towards the surface in lateral temporal areas and toward middle layers in inferior frontal areas, whereas for predictable content, we expected strong activation in deep layers of lateral temporal areas. Consistent with previous research, we found that, predictable content significantly decreased reaction times compared to unpredictable content. At the group level, the difference between valid and invalid stimuli and also the difference between valid and predictable omitted stimuli modulated superficial cortical layers in inferior frontal regions. The difference between predictable and unpredictable stimuli modulated only early auditory cortex region planum polare with no significant interaction between layer and condition. When the stimuli were omitted, the difference between predictable omitted and unpredictable omitted stimuli modulated early and intermediate auditory regions mostly in superficial layers. Using univariate tests we could not distinguish the response to the heard syllable in the cortex or hippocampus. Using decoding, we found significant above chance level decoding accuracy of the specific syllable in hippocampus area CA3 using valid trials, with increased decoding including the invalid trials coded based on the expectation but not when coding them based on heard stimulus. In sum, content validity showed a consistent modulation in middle and superficial IFG regions in line with the canonical micro-circuit for predictive coding suggesting that prediction errors are encoded in supragranular layers (Bastos et al., 2012). Stimulus predictability can be seen as a signature of precision since the predictable stimuli have 75% chance and the unpredictable stimuli 50% chance and thus an increased precision of the predictable content recruits more intermediate areas when the predictable content is omitted. In addition, our multivariate analyses support the idea that area CA3 in the hippocampus encodes stimulus expectations rather than stimulus content.

ID: 120
Abstract
Keywords: Temporal integration, vigilance, sedation, arousal, ferret

Influence of vigilance states on neuronal integration windows in the ferret auditory cortex

Hortense Gouyette[1], Magdalena Sabat[1], Arsenii Goriachenkov[1], Pierre Orhan[2], Sam V Norman-Haignere[3], Yves Boubenec[1]
[1]Ecole Normale Supérieure, France; [2]Institut du Cerveau, France; [3]University of Rochester, USA; hortense.gouyette@ens.psl.eu

Temporal integration is fundamental to auditory perception. Recent findings suggest that auditory neurons in the ferret are primarily organized along the auditory hierarchy, with longer integration windows in non-primary auditory areas (Sabat et al., 2025). This pattern aligns with human ECoG data (Norman-Haignere et al., 2022) and appears largely independent of stimulus properties.

Neural responses to repeated stimuli exhibit inherent variability, which has increasingly been studied in relation to arousal—a set of ongoing physiological processes that shape sensory processing (Tsodyks et al., 1999; Stringer et al., 2019). Pupil size is commonly used as a proxy for arousal, revealing its influence on response magnitude (McGinley et al., 2015; Saderi et al., 2021) and frequency tuning (Lin et al., 2019). Nevertheless, little is known on how variations in arousal impact temporal integration in the auditory system.

Here, we investigate whether temporal integration windows vary with arousal state and vigilance level. We recorded neural activity in the ventral primary and secondary auditory cortex of ferrets across wakefulness, REM sleep, NREM sleep, and sedation. Sleep scoring was performed with olfactory bulb gamma and hippocampal theta activity (Goriachenkov, Mahéo et al. in prep). Using microelectrode floating arrays and the temporal context invariance paradigm (Norman-Haignere et al., 2022), we directly measured integration windows in single- and multi-units, estimating them via a computational model. Additionally, we recorded physiological markers such as breathing rate, heart rate, and pupil size to assess arousal through multiple metrics.

Preliminary results suggest that integration windows remain remarkably stable across arousal states, despite substantial variations in mean firing rate. Notably, sedation appears to increase integration width. Future analyses will further compare wake, REM, and NREM stages to refine our understanding of how vigilance states influence temporal integration.

ID: 121

Abstract

Keywords: Synaptic Plasticity, E/I balance, Learning and Memory, Structural Imaging, Intrinsic Brain Dynamics

Imaging collective synaptic dynamics in the mouse auditory cortex during learning

Altug Kamacioglu[1], Pamela Osuna Vargas[2], Dominik Aschauer[1], Matthias Kaschube[2], Simon Rumpel[1]

[1]Institute of Physiology, FTN, University Medical Center, Johannes Gutenberg University Mainz; [2]Frankfurt Institute forAdvanced Studies & Department of Computer Science, Goethe University Frankfurt; akamacio@uni-mainz.de

There is strong evidence that alterations in the structural and functional properties of synaptic connections during learning are underlying the formation of long-term memories. However, how learning-induced plasticity introduces specific changes in the architecture of the neuronal network is still poorly understood. A major limiting factor of current in vivo imaging studies of synaptic dynamics is that typically only a small number of synapses are assessed in a given animal, typically tens to hundreds. Here, we densely bulk-labeled endogenous PSD-95, a major postsynaptic scaffolding protein of excitatory synapses in layers 1-3 of the auditory cortex in mice using AAV-mediated expression of FINGR intrabodies. Applying longitudinal in vivo two-photon microscopy, we re-image the same volumes of neuropil in 12 imaging sessions over an 11-day period containing more than a million synapses in an individual mouse. Using the same imaging schedule, we image synaptic dynamics in mice in control groups during environmentally and behaviorally stable conditions as well as in mice undergoing classical conditioning using auditory cues, a paradigm we have previously shown to induce a transient disbalance in the ongoing formation and elimination of dendritic spines in the auditory cortex. In parallel, we are developing an automated detection and tracking methodology for PSD-95 puncta to quantitatively analyze synaptic changes. With this dataset, we hope to provide an entry point for the study of the collective dynamics of synapses during learning, but also basal conditions.

ID: 122
Abstract
Keywords: decoding, synergy, spectrotemporal receptive field, redundancy, neural code

The reliability of neural response as a marker of the transition of neural codes along auditory pathways

Alexa Buck, Typhaine Dupont, Rupert Andrews Cavanagh, Olivier Postal, Jérôme Bourien, Nicolas Michalski, Boris Gourevitch
Université Paris Cité, Institut Pasteur, AP-HP, INSERM, CNRS, Fondation Pour l'Audition, Institut de l'Audition, IHU reConnect; Paris, France; boris@pi314.net

The detection and identification of relevant acoustic features in the environment to categorize sounds is one of the most prevalent tasks of the auditory system . Despite the variability in the presentation of the same acoustic feature due to changes in emitter's localization or timbre, or in acoustic environment, the neural representations of these varying acoustic features elicit similar perceptions suggesting the existence ofs a robust stimulus-response function between complex sounds and neural populations activity at all stages of the auditory system. We tested this hypothesis by decoding a random sound stream using spike trains from a biophysical model of the auditory nerve and from large-scale recordings in the inferior colliculus, the auditory thalamus and the auditory cortex of awake mice. Our decoding algorithm estimates the stimulus by acoustic pieces having elicited similar individual or populational neural responses in terms of either temporal firing patterns or firing rate. For individual neurons, we found that the reliability of temporal and rate codes decreased along the ascending auditory pathways while rate coding was progressively favoured with increasing independence to neuron frequency tuning. We also showed that the firing rate of auditory cortex neurons is synergistic compared to more redundant responses at subcortical stages and that this served an improved coding of sound complexity by ensembles rather than individual neurons. Finally, we demonstrated that combinatorial codes involving neural firing and neural silence within neuron pairs were efficient at coding sound information, particularly in the auditory cortex. Overall, our results picture the progressive transformation of the neural code from an individual, redundant and temporal code at the periphery to a more distributed rate-based code in the auditory cortex, based on the reliability of the neural response to variations in complex sounds.

ID: 123
Abstract
Keywords: behavior, facial motion, natural sounds, ferret, motion energy

Sound-evoked facial motion in ferrets: behavioral responses to natural and artificial sounds

Mathilde Martin, Yves Boubenec
Laboratoire des Systèmes Perceptifs, Département d'Études Cognitives, École Normale Supérieure PSL Research University, CNRS, Paris, France;
mathilde.martin@ens.psl.eu

Auditory stimuli are known to elicit facial and bodily movements in mice, reflecting underlying neural processes. In particular, sound can modulate activity in sensory regions such as the primary visual cortex, with previous studies suggesting that this modulation arises not from direct projections of the auditory cortex, but from changes in internal states and behavior. Moreover, sound-evoked facial motion has been shown to be highly sensitive in mice, even at low sound levels, and capable of tracking slow modulations in acoustic signals.

In this behavioral study, we investigated whether similar sound-induced facial movements occur in ferrets, a species for which such responses remain largely unexplored. Using high-frame rate video recordings of head-fixed ferrets passively listening to simple artificial auditory stimuli, including silent gaps in noise and descending frequency sweeps, we analyzed facial motion energy (FME) and observed a significant increase in facial motion at noise onset. FME reliably tracked the sound envelope at modulation rates up to 2 Hz, with coherence decreasing at higher sweep rates, consistent with prior mice data.

We further extended our investigation to include natural sounds from various categories (e.g., conspecific, predators and prey vocalizations, human speech, and environmental sounds) and their corresponding model-matched versions, which are synthetic sounds generated to match spectrotemporal features of the originals based on auditory perception models. We observed a significant effect of sound level on average FME, with higher levels producing more pronounced movements, and synthetic sounds producing a higher average FME. Despite notable inter-individual variability, analysis of the temporal response functions derived from the FME revealed a consistent response to sound onset, with an approximate latency of 100 ms, indicating sensitivity to low-frequency envelope fluctuations.

This work provides a foundational behavioral characterization of sound-evoked facial motion in ferrets, with the dual aim of identifying stimuli that most robustly elicit motion for future investigations and offering preliminary insights into the relationship between sensory-driven movement and neural activity.

ID: 124

Abstract

Keywords: Autism Spectrum Disorder, Hypersensitivity, Auditory Cortex, Calcium Imaging, Behavior

Altered Auditory Cortex Neuronal Dynamics in ASD modeled mice: Implications for Auditory Hypersensitivity

Netta Baram[1], Hila Sapir[1], Ghattas Bisharat[1], Jennifer Resnik[1,2]

[1]Department of Life Sciences, Ben Gurion University of the Negev, 84105, Beer Sheva, Israel; [2]Zelman Center for Brain Science Research, Ben-Gurion University of the Negev, 84105, Beer Sheva, Israel; bnett@post.bgu.ac.il

Autism spectrum disorder (ASD) is frequently associated with atypical sensory processing, with auditory hypersensitivity being one of the most common sensory-related symptoms. To investigate the neural mechanisms underlying altered sound perception in ASD, we combined in vivo two-photon calcium imaging with behavioral assays in a rodent model. Imaging of auditory cortical activity in ASD mice (n=5) revealed increased spontaneous and sound-evoked activity, along with a higher percentage of sound-responsive neurons compared to WT controls (n=4). To assess whether these cortical differences impact perception, we trained mice on a two-alternative forced-choice (2AFC) loudness discrimination task. ASD mice were more likely to categorize intermediate-intensity sounds (44–65 dB SPL), perceived as soft by WT mice, as loud. These findings suggest that increased auditory cortical excitability may contribute to heightened loudness perception in ASD, providing insight into the neural basis of auditory hypersensitivity in the disorder.

ID: 126
Abstract
Keywords: auditory cortex, perception, magnetoencephalography

Laminar specificity of the auditory perceptual awareness negativity: a biophysical modeling study

Carolina Fernandez Pujol[1], Elizabeth Blundon[2], Andrew R. Dykstra[3]
[1]Brown University, Providence, RI, United States of America; [2]Nova Scotia Health, Halifax, NS, Canada; [3]University of Central Florida, Orlando, FL, United States of America; andrew.dykstra@ucf.edu

How perception of sensory stimuli emerges from brain activity is a fundamental question of neuroscience. To date, two disparate lines of research have examined this question. On one hand, human neuroimaging studies have helped us understand the large-scale brain dynamics of perception. On the other hand, work in animal models (mice, typically) has led to fundamental insight into the micro-scale neural circuits underlying perception. However, translating such fundamental insight from animal models to humans has been challenging. Here, using biophysical modeling, we show that the auditory awareness negativity (AAN), an evoked response associated with perception of target sounds in noise, can be accounted for by synaptic input to the supragranular layers of auditory cortex (AC) that is present when target sounds are heard but absent when they are missed. This additional input likely arises from cortico-cortical feedback and/or non-lemniscal thalamic projections and targets the apical dendrites of layer-5 (L5) pyramidal neurons. In turn, this leads to increased local field potential activity, increased spiking activity in L5 pyramidal neurons, and the AAN. The results are consistent with current cellular models of conscious processing and help bridge the gap between the macro and micro levels of perception-related brain activity.

ID: 127
Abstract
Keywords: Primary auditory cortex, fMRI, frequency discrimination, hyperacuity, cortical magnification

Does frequency magnification in primary auditory cortex predict individual frequency hyperacuity thresholds?

Ričards Kazibrodskis[1], Moussa Kousa[2,3], Issam Tafech[2,4,5], Omar Ibrahim[2,4], Julien Besle[1,2]
[1]University of Plymouth, United Kingdom; [2]American University of Beirut, Lebanon; [3]The Pennsylvania State University, United States; [4]Charité – Universitätsmedizin Berlin, Germany; [5]Humboldt-Universität zu Berlin, Germany; julien.besle@plymouth.ac.uk

The auditory system over-represents (or magnifies) low frequencies compared to high frequencies, both in the cochlea and primary auditory cortex. Previously, we showed that the cortical frequency magnification (CFM) function (the relative surface area dedicated to different frequencies along a tonotopic map), measured using fMRI and averaged across listeners, is more closely related to published estimates of frequency discrimination (greater CFM predicts a smaller discrimination thresholds) than to published estimates of cochlear frequency selectivity (Gurer et al., 2023). This suggests that frequency discrimination, a form of frequency hyperacuity, is primarily constrained by cortical rather than cochlear processes. Here we attempt to improve on this by measuring the CFM and frequency discrimination functions in the same participants, and test whether CFM predicts individual discrimination thresholds.

We measured pure tone difference limens for frequency (DLFs) in 18 normal-hearing listeners at 8 log-spaced frequencies between 0.2 and 8 kHz, using an updated maximum likelihood adaptive procedure. To measure the CFM function, we measured the BOLD-fMRI response (3T, sparse GE-EPI, TR = 8s, 1.5 mm resolution) to 4.5-second trains of pure tones around the same 8 frequencies. We then estimated preferred frequency and frequency tuning width at each cortical surface vertex by fitting log-frequency Gaussians to the BOLD frequency tuning curves. All participants showed two mirror-reversed tonotopic gradients separated by a low-frequency reversal along Heschl's gyrus, in a region of elevated frequency selectivity corresponding to primary auditory cortex. Finally, we derived the CFM function in each tonotopic gradient from 3 alternative measurements: (1) the cortical position of each vertex along the tonotopic gradient, (2) the proportion of vertices responding to different frequencies and (3) the local preferred frequency gradient strength, each plotted as a function of vertices' preferred frequency.

We fitted the DLF function and the reciprocal of the CFM function (averaged across hemispheres and primary tonotopic gradients) of each participant with the same DLF equation (Micheyl et al., 2012) and estimated participant-specific linear parameters representing overall DLF/CFM (intercept) and change in DLF/CFM with frequency (slope). Preliminary results with a subset of 12 participants suggest no correlation between DLF and CFM parameters, for either intercept or slope, using either of the 3 CFM measurement methods. This unexpected lack of correlation could be due to averaging across frequencies, low reliability of DLF/CFM estimates at low and high frequencies, or to an imperfect fit of the DLF equation to the empirical CFM function (Gurer et al., 2023).

Gurer et al., 2023. Frequency mapping in human primary auditory cortex predicts frequency discrimination performance. Presented at ISAAR, Nyborg.

Micheyl et al., 2012. Hearing Research 292, 1–13.

ID: 128

Abstract

Keywords: speech, timescale, intracranial EEG, ECoG, computational models

Temporal integration in human auditory cortex is predominantly yoked to absolute time, not structure duration

Sam Norman-Haignere[1], Menoua Keshishian[2], Orrin Devinsky[3], Werner Doyle[3], Guy McKhann[4], Catherine Schevon[4], Adeen Flinker[3], Nima Mesgarani[2]

[1]University of Rochester Medical Center, United States of America; [2]Columbia University, United States of America; [3]NYU Langone Medical Center, United States of America; [4]Columbia University Medical Center, United States of America; samuel_norman-haignere@urmc.rochester.edu

The sound structures that convey meaning in speech such as phonemes and words vary widely in their duration. As a consequence, integrating across absolute time (e.g., 100 ms) and sound structure (e.g., phonemes) reflect fundamentally distinct neural computations. Auditory and cognitive models have often cast neural integration in terms of time and structure, respectively, but whether neural computations in the human auditory cortex reflect time or structure remains unknown. To answer this question, we rescaled the duration of all speech structures using time stretching/compression. We then measured neural integration windows for stretched and compressed speech using a recently developed experimental paradigm and computational model (the temporal context invariance or TCI paradigm). The TCI paradigm estimates the shortest sound segment for which the neural response is invariant to surrounding context stimuli. The paradigm does not make any assumptions about the stimulus features that underlie the neural response (e.g., a spectrogram) or the mapping between those features and the neural response (e.g., linear) and thus is effective in highly nonlinear systems like the brain. We show that our approach is effective at distinguishing time- vs. structure-yoked integration from computational models, including revealing a transition from time- to structure-yoked computation across the layers of a popular deep neural network model trained to recognize structure from natural speech. To determine if neural computations in the human auditory cortex are yoked to time or structure, we applied our paradigm to spatiotemporally precise intracranial recordings from neurosurgical patients. We observed that integration windows were significantly longer for stretched vs. compressed speech, but this lengthening was very small (~5%) relative to the change in structure durations, even in non-primary regions strongly implicated in speech-specific processing. Similar results were observed for speech that was naturally spoken at a fast or slow rate rather than being uniformly stretched or compressed. These findings demonstrate that time-yoked computations dominate throughout the human auditory cortex, placing strong constraints on neurocomputational models of structure processing.

ID: 129
Abstract
Keywords: Cat, Vocalizations, Acoustic behaviour, Temporal cortex

A HIERARCHICALLY ORGANIZED SOUND DISCRIMINATION PATHWAY IN AUDITORY CORTEX

Stephen G. Lomber[1,2], Paul Cornwell[3]
[1]University of Texas at Arlington, USA; [2]McGill University, Canada; [3]Pennsylvania State University, USA; steve.lomber@mcgill.ca

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

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

ID: 131

Abstract

Keywords: auditory sequence processing, structure learning, random network, EEG

Higher-Order Network Structures mediate Neural Responses to Tone Sequences

Xiaoyue Zhou[1,4], Muzhi Wang[1,2,3], Ying Fan[1,2,3], Huan Luo[1,2,3]

[1]School of Psychological and Cognitive Siences, Peking University, Beijing, China; [2]PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China; [3]Beijing Key Laboratory of Behaviour and Mental Health, Peking University, Beijing, China; [4]Ear Institute, University College London, United Kingdom; xiaoyue.zhou@ucl.ac.uk

Background

Music, as a unique human experience, is widely enjoyed and remembered. When memorizing music, we do not merely recall isolated tones but also extract and retain the hierarchical hidden structures embedded within these tones. Understanding how humans extract and represent such hidden structures provides deeper insights into how we abstract knowledge from sensory experiences, which in turn facilitates memory encoding and retrieval.

Methods

Conceptualizing transitions between musical tones as a network, we generated pure tone sequences based on a random network to serve as experimental materials. We first exposure participants with a 3-min sequence, then instructed them to complete a sequential prediction task, where they needed to report whether the last tone in a short 6-tone sequence sounded expected. The study included a training phase (day1) and a test phase (day2, with EEG recordings). Higher-order structural information is measured using the communicability between tones.

Results

Behavioral analyses revealed significant modulations of responses and reaction times (RTs) by the underlying structure: (1) The likelihood of reporting the final tone as expected increased with the minimal distance between the last two tones; (2) RTs decreased with higher communicability when the final tone was reported as expected, but increased when the final tone was reported as unexpected.

EEG analyses further demonstrated neural representations of higher-order sequence structures: (1) The amplitude of Mismatch Negativity (MMN) scaled with communicability and predicted behavioral responses; (2) ERPs to tones continuously tracked communicability, even when transition distances were controlled; (3) the more ERP can be explained by communicability, the more accurate participants made behavioral judgement.

Summary

Humans learn and predict auditory sequences not only by remembering local connection patterns, but also by extracting transitional network structures between tones. This process further enables efficient information processing based on the computation of higher-order structure properties.

ID: 132

Abstract

Keywords: auditory cortex, spectrotemporal receptive field, two-photon calcium imaging

Spectrotemporal receptive fields of auditory cortical responses measured with 2-photon calcium imaging

Arev Aksel Melikyan, Johannes Dahmen, Ben Willmore, Andrew King, Kerry Walker

Department of Physiology, Anatomy and Genetics, University of Oxford, United Kingdom; arev.melikyan@exeter.ox.ac.uk

Neurons in auditory cortex integrate sound stimuli over time and across frequency to represent complex features. The spectrotemporal receptive field (STRF) is a commonly used encoding model which can efficiently derive a linear response filter of an auditory neuron from its eletrophysiologically measured spiking response to complex sounds. Two-photon calcium imaging allows us to sample the activity of a larger number of single neurons across wider cortical areas than microelectrode recordings, but lacks the temporal precision to resolve individual spikes. Therefore, it is unclear if STRFs can be reliably applied to calcium imaging data. To address this question, we fit STRFs to neural responses obtained from in vivo two-photon calcium imaging in mouse auditory cortex. Preliminary data indicate that receptive fields derived from imaging experiments can have spectrotemporal structure and predictive power comparable to that of electrophysiologically-derived STRFs when parameters are optimized for calcium data. Furthermore, we use a spike-to-GCaMP-fluorescence forward model to evaluate the effect of this transformation on measured STRFs. The results will allow future studies to investigate the topography of spectrotemporal tuning at both local and global scales in auditory cortex with 2-photon imaging, with the benefits of genetically targeting cell types and pathways.

ID: 133
Abstract
Keywords: Speech perception, SEEG, auditory cortex, theta, gamma

Theta and gamma rhythms in the human auditory cortex are evoked by speech

Victor J. Lopez-Madrona[1,2], Jeremy Giroud[3], Manuel Mercier[1], Bruno Giordano[4], Agnes Trebuchon[1,5], David Poeppel[6], Benjamin Morillon[1]
[1]Aix-Marseille Université, INSERM, INS, Institut de Neurosciences des Systèmes, Marseille, France, Japan; [2]Institute of Language, Communication, and the Brain, Aix-Marseille Univ, Marseille, France; [3]MRC Cognition and Brain Sciences Unit, University of Cambridge, UK; [4]Aix Marseille Univ, CNRS, INT, Marseille, France; [5]APHM, Timone Hospital, Epileptology and Cerebral Rhythmology Department, Marseille, France; [6]New York University, 6 Washington Place, New York, NY, USA; victor.lopez-madrona@univ-amu.fr

Theta and gamma oscillations dominate the auditory cortex during speech perception, but their origin remains controversial. Using intracerebral recordings from 18 epilepsy patients, we show that speech acoustics shape these oscillations, rather than reflecting endogenous computations. At rest, alpha oscillations dominate the auditory cortex and induced neural dynamics during speech perception are restricted to alpha-beta (5–20 Hz) suppression and high-frequency (80–120 Hz) enhancement. Instead, theta and gamma neural responses are linearly driven by speech, reflecting the syllabic rate (2–6 Hz), vowel features (30–50 Hz), and the fundamental frequency (100–150 Hz). Neural theta-gamma coupling also originates from the speech stimulus and is a consistent acoustic signature of speech across 17 languages. Finally, neural gamma amplitude modulates theta phase, reflecting faster coupling of higher neural frequencies with acoustic input. These results refute the hypothesis that theta-gamma dynamics reflect endogenous computations in speech perception, showing instead that they originate from the speech signal itself, calling into question their functional role in speech processing.

ID: 134

Abstract

Keywords: otoferlin, congenital deafness, gene therapy, cochlear nucleus, inferior colliculus

Effects of gene therapy on central auditory processing in a mouse model of congenital DFNB9 deafness

Ara Schorscher-Petcu, Mauricio Saenz, Najate Benamer, Saaid Safieddine, Jérémie Barral

IHU reConnect, Institut Pasteur, France; ara.schorscher-petcu@pasteur.fr

Normal hearing relies on the fast and precise transmission of information between sensory inner hair cells (IHCs) and afferent auditory neurons owing to the highly specialized ribbon synapse connecting the two. Otoferlin is a calcium sensor essential for ultra-fast and sustained transmitter release in IHCs.

Mutations in otoferlin encoding gene, OTOF, cause profound deafness known as DFNB9. Using an Otof KO mouse as a model, we have previously developed gene therapy with murine Otof cDNA and successfully restored hearing in congenitally deaf mice. The human version of this gene therapy, using human OTOF cDNA is currently in phase II clinical trials.

Here, we compare the efficiency of murine vs human cDNA sequences to restore hearing in a DFNB9 mouse model. We find that although both coding sequences reinstate expression and subcellular localization of otoferlin, the human sequence performs less well in restoring hearing as measured by auditory brainstem responses (ABR).

To investigate how differential hearing restoration affects central sound processing, we used high-density probes to record neuronal responses simultaneously from the cochlear nucleus (CN) and the inferior colliculus (IC) in wild-type and KO mice treated with murine or human cDNA. We find that responses evoked by amplitude modulated tones in the CN of KO mice treated with murine cDNA resemble those of wild-type mice. In contrast, discharge rates after treatment with human cDNA are significantly lower. Neural responses in the IC of the human cDNA group are also reduced, but less so and with the late component being more affected. In addition, stimulus-evoked spike-latencies are longer in both structures and entrainment is lost beyond 100 Hz in the human cDNA group. In line with this, stimulus phase-locking is reduced in individual CN units, as is synchronization at the population level in both structures after treatment with human cDNA. These results, both at the level of ABRs and single neuron data, could be reproduced by altering the kinetics of synaptic fusion in a biophysical model of the auditory system. Finally, we show that the deficits pervading after treatment with human cDNA perturb frequency encoding in CN, but not in IC, pointing towards compensatory mechanism in the latter.

This work will be relevant to optimizing gene therapy for congenital deafness and to understanding how the central nervous system adapts to partial or complete post-developmental restoration of sensory input.

ID: 135
Abstract
Keywords: Temporal envelope modulations, ASSR, EEG, Specialization, Psychoacoustics

Investigating the relation between temporal envelope modulations and speech perception

Brent Sterckx[1,2], Marc Moonen[1], Jan Wouters[2]
[1]STADIUS, Dep. of Electrical Engineering, KU Leuven, Belgium; [2]ExpORL, Dep. of Neurosciences, KU Leuven, Belgium; brent.sterckx@kuleuven.be

Natural speech is a complex acoustical signal that contains both spectral and temporal information. Temporal information can be subdivided into periodicity, temporal fine structure (TFS) and temporal envelope modulations (TEMs). Among these, especially TEMs are crucial for speech perception as they encode important temporal cues at several different time-scales, i.e. ranges of modulation frequencies. For instance, the speech typically contain TEMs at ~4 Hz (250 ms) and ~20 Hz (50 ms), corresponding to resp. the syllable and phoneme rates. Recent neurophysiological findings suggest a hemisphere-specific specialization for processing distinct ranges of modulation frequencies. Coupled with the typical contralateral dominance of auditory pathways, this specialization may underlie the right-ear advantage observed in speech perception tests. However, it remains unclear how overall speech perception can be decomposed into the contributions due to the different components of the natural speech signal in each ear.

In contrast to this top-down approach – dissecting natural speech to disentangle overall performance – a bottom-up approach instead selectively modulates a synthetic carrier at specific modulation frequencies to mimic the basic building blocks of speech. This stimulus evokes an auditory steady-state response (ASSR) that is often presented as an objective measure probing the brains TEM processing at that modulation frequency. Yet, it has so far not been shown that the results obtained using these simplistic stimuli are specifically indicative of the contribution of the encompassing range of TEM frequencies of the speech signal to the overall performance, or whether they rather reflect more general TEM processing.

The goal of this study was therefore to decrease the gap between such two approaches to explaining TEM-based speech perception, by on the one hand separating the overall speech perception performance with natural speech into the contributions from each component, and on the other hand to investigate the connection between these contributions and corresponding ASSRs.

In the top-down approach, natural speech was processed to selectively retain only specific modulation frequency ranges and periodicity and TFS information. This allowed to measure isolated contributions to the speech perception performance due to the retained components only, disentangling the overall performance. In the bottom-up approach, modulated noise stimuli were used to elicit ASSRs at modulation frequencies matching the top-down approach.

Together, these complementary approaches aim to provide a more comprehensive understanding of the role of specific TEMs and side of stimulation in speech perception, and to assess the precision of ASSRs as biomarkers for TEM processing.

Acknowledgements: This work was funded through a fellowship strategic basic research from the FWO (grant no. 1SHI924N), awarded to B.S. The authors thank the participants for donating their time.

ID: 136

Abstract

Keywords: hearing loss, dementia, vascular plasticity, congenital deafness, hearing restoration

Vascular plasticity of hte auditory cortex in health and disease

Sophie Skriabine[1,3], Elisa De Launoit[1,3], Sabrina Méchaussier[2], Etienne Pérès[2], Philippe Jean[2], Typhaine Dupont[2], Rémi Pommé[1], Domitille Rajot[1], Etienne Doumazane[1], Charly Rousseau[1], Marie-José Lecomte[2], Ghizlene Lahlou[2], Saaid Safieddine[2], Nicolas Michalski[2,4], Nicolas Renier[1,4]

[1]Laboratory of Structural Plasticity, Sorbonne Université, ICM Paris Brain Institute, INSERM U1127, CNRS UMR7225, 75013 Paris, France; [2]Université Paris Cité, Institut Pasteur, AP-HP, INSERM, CNRS, Fondation Pour l'Audition, Institut de l'Audition, IHU reConnect, Paris, F-75012, France.; [3]These authors contributed equally, listed anti-alphabetically; [4]Co-senior authors; nicolas.michalski@pasteur.fr

Age-related hearing loss was recently shown to be a major risk factor for dementia. Hearing loss deeply modifies the neuronal activity and connectivity of auditory networks and their associated regions, but less attention has been given to its effects on the brain vascular network, which has been implicated in all forms of dementia. The need for speed and temporal accuracy of the auditory system results in a high metabolic demand. A correlate of this need is that the auditory system might be highly vulnerable to changes in oxygen and glucose supply and is probably sensitive to the "health" of the vascular network. We combined recent methodological progress in reconstructing the 3D organization of the vasculature at the scale of the entire brain and mouse models of deafness mimicking different types of auditory deficits, and hearing restoration. We show that hearing impairment affects the vascular organization throughout the brain. Congenital deafness, (Otoferlin knock-out mouse model), caused a decrease of capillary density in all auditory regions throughout the brain, starting at 2 months of age. Surprisingly, these deficits expanded to many other non-auditory cortical regions, and worsened for many of them with aging, suggesting that a capillary deficit in auditory cortical regions may propagate to other cortical areas. Finally, hearing restoration using the Otof gene therapy at one month of age stabilized the vascular network and prevented further impairment of the vascular network. Our work establishes that hearing loss is detrimental to the vascular network and addresses a long-standing controversy regarding the potential of activity-dependent adult vascular structural plasticity. By determining whether cerebrovascular deficits can be reversed by hearing restoration, this work provides a scientific basis for improving auditory rehabilitation in patients and for evaluating the use of hearing aids to decrease the risk of dementia in the general population suffering from age-related hearing loss.

ID: 137

Abstract

Keywords: Auditory cortex; Deep neural network; Speech perception; Music perception

Deep Sound Synthesis Reveals Novel Category-Defining Sound Features in the Human Auditory Cortex

Lidongsheng Xing, Elia Formisano, Lars Riecke

Maastricht university, Netherlands; l.xing@maastrichtuniversity.nl

The human auditory system extracts meaning from the environment by transforming acoustic input signals into semantic categories. Specific acoustic features give rise to distinct categorical percepts, such as speech or music, and to spatially distinct preferential responses in the auditory cortex. These responses contain category-relevant information, yet their representational level and role within the acoustic-to-semantic transformation process remain unclear. We combined neuroimaging, a deep neural network, and a brain-based sound synthesis to identify the sound features that are internally represented in the speech- and music-selective human auditory cortex. In subsequent in-silico, neuroimaging, and behavioral experiments, we tested the functional role of these features for sound categorization. We found that the synthetized sounds exhibit unnatural features distinct from those normally associated with speech and music, yet they elicit categorical cortical and behavioral responses resembling those of natural speech and music. The identified categorization-relevant features differ from those promoted by common acoustic models (e.g. spectrotemporal modulation or texture models). Our findings provide novel insights into the fundamental sound features underlying speech and music categorization in the human auditory cortex.

ID: 138
Abstract
Keywords: speech perception, EEG, auditory processing, speech intelligibility

Tracking of Acoustic and Linguistic Speech Features by the Brain

Alexis Deighton MacIntyre, Tobias Goehring, Matt Davis
MRC Cognition and Brain Sciences Unit, University of Cambridge, United Kingdom; alexisdeighton.macintyre@mrc-cbu.cam.ac.uk

During speech listening, patterns of neural activity become temporally coupled to stimulus features, which can be measured using electroencephalography (EEG) and other neuroimaging techniques. This opportunity to "decode" from the listening brain may provide important insights into the nature of speech perception. It also holds promise as a candidate objective measure to guide fitting and assess outcomes for cochlear implant (CI) listeners. The effects of CI-associated spectral degradation on neural speech decoding, however, remain unclear. In particular, the interplay between speech intelligibility and acoustic signal quality is debated, with contradictory evidence for and against the enhancement of neural speech tracking for comprehended speech. Importantly, few studies independently manipulate acoustic properties and intelligibility. In the current experiment, we simulate CI listening by presenting natural and spectrally degraded speech to typically hearing listeners (n = 38) undergoing EEG recording. To dissociate acoustic-sensory processing from comprehension, we use intelligible (English) and non-intelligible (Dutch) speech produced by one bilingual speaker. We promote auditory attention throughout the experiment using a novel prosodic target detection task that is not contingent on speech understanding. Decoding models were then trained to reconstruct speech parameters from held-out neural response data, with correlations between reconstructed and true stimulus features providing a measure of neural speech decoding. We discuss our first analysis of this data set, which focused on the speech amplitude envelope, a commonly used acoustic feature within the speech decoding literature. Turning to ongoing work, we share early results concerning how well other, less frequently explored acoustic properties of the speech signal, such as spectral flux, can be reconstructed from EEG. Finally, the decoding of linguistic features, such as word and phoneme onsets, are compared to speech acoustics with a view towards identifying the parameters that best indicate differences in sensory-cognitive speech processing across listening conditions.

ID: 139
Abstract
Keywords: Auditory neuroscience, large language models, modeling, auditory attention

Large Language Models Embeddings as Predictors of Electrocorticography Responses in a Multi-Talker Attention Paradigm

Corentin Puffay[1,3], Gavin Mischler Moore[2], Vishal Choudhari[2], Jonas Vanthornhout[1], Richard Antonello[2], Sukru Samet Dindar[2], Hugo Van hamme[3], Tom Francart[1], Nima Mesgarani[2]

[1]KU Leuven, Dept. Neurosciences, ExpORL, Leuven, Belgium; [2]Columbia University, Department of Electrical Engineering, New York, NY, United States; [3]KU Leuven, Dept. of Electrical engineering (ESAT), PSI, Leuven, Belgium; corentin.puffay@kuleuven.be

This study investigates how auditory attention modulates context tracking in the brain in a multi-talker environment. Some studies demonstrated the efficiency of large language models (LLMs) in modeling contextualized semantic encoding in the human brain. The transcripts of the presented speech are fed to an LLM to generate a contextualized representation of words used as speech features, further used in linear regression models to predict the brain response. However, very few studies investigated their modeling performance in more challenging scenarios. We explore how LLM embeddings can predict brain responses to continuous speech in a two-conversation "cocktail party" paradigm. Using electrocorticography (ECoG) and stereoelectroencephalography (sEEG) data from three epilepsy patients, we compared prediction scores between different LLM layers for both attended and unattended speech streams.

Our results show that LLMs generate language features that models can effectively use to predict brain responses to the attended conversation. Still, surprisingly, contextual information from the unattended stream also carries predictive power. Furthermore, our findings suggest that the brain tracks the context on a faster time scale for the unattended stream compared to the attended one. These results provide new insights into the complex role of attention in context encoding, suggesting that the brain processes multiple streams of speech simultaneously, albeit with differing contextual temporal scopes. In addition, these results emphasize the potential of LLMs as tools for studying neural language processing in complex auditory environments.

ID: 140
Abstract
Keywords: vocalization, marmoset, non-human primate, vocal production, error signal

Role of Auditory Cortex in Vocal Control

Steven J Eliades[1], Joji Tsunada[2]
[1]Duke University School of Medicine, United States of America; [2]Chinese Institute for Brain Research, China; steven.eliades@duke.edu

Vocalization is a sensory-motor process requiring auditory self-monitoring to detect and correct errors in vocal production. Recent evidence has suggested that the auditory cortex is involved in coding vocal feedback errors through a predictive vocalization-induced suppression. Supporting a potential role in feedback-dependent vocal control, error-related activity in auditory cortex has been found to predict subsequent vocal acoustics, and electrical microstimulation has been shown to evoke rapid changes in vocal production. Key questions remain, however, as to whether or not auditory error responses predict specific compensatory vocal behaviors, or change with varying feedback conditions, and whether auditory cortex is even necessary for feedback vocal control. In this study, we investigated neural responses in the auditory cortex of vocalizing marmoset monkeys, testing frequency-shifted feedback of varying direction. We found that neural activity correlated with subsequent vocal acoustic changes, and that this activity consistently predicted the direction of feedback compensation, with most units predicting compensation for one feedback direction and null responses for feedback in the other; units switching the direction of their vocal behavioral predictions were rare. In order to examine a potential causal role for this activity, we performed muscimol injections into the auditory cortex, and measuring effects on acoustic changes during altered vocal feedback. In addition to changes in overall vocal acoustics, we found that muscimol injections reduced the magnitude of vocal feedback compensation, when compared to sham injections at the same sites. Therse results suggest that vocal error-related neural activity in the auditory cortex drives a consistent behavioral response, possibly due to specific mapping onto vocal motor control circuits, and that engaging specific neural sub-populations may drive dynamic control of vocal frequency in a 'push-pull' fashion. Importantly, these results also suggest that cortical activity may play a causal role in our ability to maintain on-line vocal stability by detecting and compensating for changes in feedback of our voice. More broadly, these findings expand our understanding of the roles played by the auditory cortex in both sensory-guided motor control and other audio-motor behaviors.

This work was supported by NIH/NIDCD Grant K08-DC014299, R01-DC018525

ID: 141
Abstract
Keywords: Speech processing, audiovisual integration, neural tracking

Speech-in-noise enhancement through natural and simplified facial movements is linked to the cortical tracking of speech in the delta band

Mike Thornton[1], Enrico Varano[2], Dorothea Kolossa[3], Steffen Zeiler[3], Tobias Reichenbach[1]
[1]FAU Erlangen-Nürnberg, Germany; [2]Universität Zürich, Switzerland; [3]Technische Universität Berlin, Germany; mike.d.thornton@fau.de

Introduction: Speech comprehension in noisy environments is enhanced when listeners can see the talker's face. Some specific visual features, such as lip movements, have been linked to audiovisual speech processing [1,2]. However, whether simplified face-like stimuli—lacking fine detail but preserving motion—support audiovisual integration remains unknown.

Methods: Nineteen normal-hearing, native English-speaking participants participated in a behavioural experiment and an electroencephalography (EEG) study.

In the behavioural study, sentences taken from the GRID corpus were presented at a signal-to-noise ratio (SNR) of -9 dB [3]. The sentences were were presented unimodally, or else they were paired with one of four different visual signals: a natural video of the talker's face; and edge-filtered version of the natural video; a cartoon derived from the natural video; or an animated disk oscillating with the temporal speech envelope. The participants completed a keyword identification task after hearing each sentence.

EEG measurements were obtained from the same participants whilst they listened to extended narratives from the AVBook corpus [4]. Speech was presented in background noise at an SNR of -2 dB. The speech material was either presented unimodally, or in combination with one of the aforementioned visual signals. We also presented the visual signals unimodally. The EEG responses to the speech envelope were characterised with encoding models, in the delta (1–4 Hz) and theta (4–8 Hz) frequency bands.

Results: Our results show that face-like stimuli enhance speech comprehension, even when fine facial details were removed, whereas simple animations of the speech envelope do not. In the EEG study, we found that the audiovisual responses to speech closesly resembled the audio-only response. However, audiovisual responses were enhanced in the delta band at a latency of 180 ms, particularly when the most natural visual signals were presented. The strength of the neural enhancement was correlated with the audiovisual gain which we measured in the behavioural study, demonstrating that neural envelope tracking in auditory cortex indexes speech-in-noise comprehension during audiovisual speech perception.

References

[1] P. Reisinger et al., "Neural Speech Tracking Contribution of Lip Movements Predicts Behavioral Deterioration When the Speaker's Mouth Is Occluded," eneuro, vol.12, no. 2. 2025.

[1] Q. Summerfield, "Use of Visual Information for Phonetic Perception," Phonetica, vol. 36, no. 4–5. 1979.

[1] M. Cooke et al., "An audio-visual corpus for speech perception and automatic speech recognition," JASA, vol. 120, no.5. 2006.

[1] E. Varano et al., “AVbook, a high-frame-rate corpus of narrative audiovisual speech for investigating multimodal speech perception,” JASA, vol. 153, no. 5. 2023.

ID: 142
Abstract
Keywords: Tinnitus, GABA, PET

GABA receptor binding in tinnitus: a PET study in humans

Marc Thioux, Emile de Kleine, Sonja J. Pyott, Antoon Willemsen, Erik de Vries, Pim van Dijk
UMC Groningen, The Netherlands; e.de.kleine@umcg.nl

Gamma-aminobutyric acid (GABA) is the main inhibitory neurotransmitter in the brain. Experiments in animals suggest that deficient GABA neurotransmission may lead to tinnitus. In addition, magnetic resonance (MR) spectroscopy studies in humans showed indications of a reduced concentration of GABA in the auditory cortex of patients with tinnitus. Together, these observations suggest a role of GABAergic neurotransmission in tinnitus.

In order to further investigate this, we performed a dynamic position emission tomography (PET) study using the 11C-Flumazenil PET tracer. Flumazenil selectively binds to GABA(A) receptors. The PET scans provide an estimate of the binding potential of flumazenil, which is a function of the availability of GABA receptors in the brain and their binding properties. PET scans were analyzed using the simplified reference tissue model (SRTM). In addition, we performed GABA-edited MR spectroscopy in the same subjects, measuring the concentration of GABA. For this, we defined two voxels encompassing the left and right primary auditory cortex, and a control voxel in the calcarine cortex.

Analysis of our PET data showed a significantly higher binding potential in the primary auditory cortex of participants with tinnitus, as compared to participants without tinnitus. Remarkably, other brain regions involved in tinnitus, did not show significant group differences. Also, MR spectroscopy in the same subjects did not show significant group differences in GABA concentration.

Consequently, GABAergic inhibition must differ between participants with and without tinnitus. These results provide strong clinical evidence for a key role of GABA in humans with tinnitus. They motivate further investigation into pharmaceutical interventions for tinnitus targeting GABAergic neurotransmission.

ID: 144

Abstract

Keywords: fMRI, implicit memory, auditory, statistical learning

Neural correlates of implicit memory for repeated regular sound patterns

Ylka Kolken[1], Chengran Li[1], Ryszard Auksztulewicz[1], Peter de Weerd[1], Vincent van de Ven[1], Roberta Bianco[2], Maria Chait[3], Federico de Martino[1]

[1]Maastricht University, Netherlands, The; [2]Italian Institute of Technology, Italy; [3]University College London, United Kingdom; ylka.kolken@maastrichtuniversity.nl

The human brain exhibits remarkable sensitivity to sensory patterns, a process documented across modalities, including vision (Saffran et al., 1999; Turk-Browne et al., 2009) and audition (Agus et al., 2010; Barascud et al., 2016; Heilbron & Chait, 2018; Paavilainen, 2013). This ability to rapidly and automatically detect sensory regularities plays a fundamental role in perceptual inference, often operating outside conscious awareness. Listeners, for example, can track emerging regularities in complex sound sequences, even when they are not behaviorally relevant (Barascud et al., 2016) and that regularity detection relies on a network including the auditory cortex, inferior frontal gyrus (IFG), and hippocampus. Repeating the same sensory pattern can lead to (implicit) memory formation, exemplified by the response time (RT) advantage for repeated patterns even after a few repetitions (Bianco et al., 2020). Using the same paradigm, we aimed to explore the neural mechanisms underlying memory formation for repeated regular patterns, using ultra-high field functional MRI (fMRI).

Participants underwent anatomical and functional MRI using a 7-T Siemens MAGNETOM scanner. They listened to 5-second sequences of 50-ms tone pips (222–2000 Hz), comprising random sequences (RAN), that either continued to be random (RANRAN), or transitioned to a regular pattern (RANREG). Prior to scanning, participants performed a 1-hour behavioural session in which they were asked to detect transitions to regularities within the auditory sequences. Unbeknownst to them, five RANREG sequences were repeated six times per block (RANREGr). In the fMRI session, these sequences were presented continuously over six runs, without a behavioural task involved.

The behavioral results showed a significant RT advantage for trained regularities (RANREGr) compared to novel regularities (RANREG) (in line with previous results - Bianco et al., 2020). Both novel and trained regularities activated the posterior and anterior STG, IFG and middle frontal gyrus (MFG). The IFG and posterior STG showed heightened activation for trained regularities, while the hippocampus exhibited reduced activity for the previously encountered patterns versus novel regularities and random sequences. These findings suggest that repeated exposure enhances neural sensitivity in auditory and memory-related regions, supporting implicit learning and retention of complex sound patterns.

ID: 145

Abstract

Keywords: ERP, LAER, auditory development, puberty, audiovisual integration

Does puberty influence auditory and audiovisual processing of speech?

Ellen Saliën, Ugo Benrubi, Axelle Calcus

ULB, Belgium; ellen.salien@ulb.be

Late Auditory Event-Related Potentials (LAERs) are widely used to examine central auditory processing, reflecting successive processing stages and being influenced by (in)congruent visual information. Across the lifespan, LAERs show morphological and topographical changes, following a stepwise maturation pattern with distinct transitions at adolescence. Puberty, marked by hormonal changes, may enhance neural plasticity in auditory regions and those critical for audiovisual integration, such as the superior temporal sulcus (STS) and temporoparietal junction (TPJ). Despite long-standing hypotheses, the role of puberty in central auditory development remains unexplored.

This study investigates pubertal effects on auditory and audiovisual speech processing. We hypothesize that (1) pubertal maturation underlies the stepwise trajectory in LAERs and (2) puberty influences visual cue impact on LAER development. Participants (n = 137; 8–21 years) were categorized into five pubertal stages (pre-, early-, mid-, late-, and post-pubertal) using the Pubertal Development Scale (PDS), alongside 10 adults (25–28 years). Stimuli (/ba/ and /da/) were presented in an oddball paradigm across four conditions: audio-only, visual-only, audiovisual congruent, and audiovisual incongruent. Saliva samples measured testosterone, estrogen, and DHEA to explore hormonal influences on LAERs.

Data collection has recently concluded, with analyses ongoing. Preliminary findings confirm a stepwise LAER maturation trajectory, aligning with prior research. A significant effect of pubertal stage emerged, especially between pre/early and mid/late phases, possibly linked to adolescent hormonal waves. However, initial N1-P2 analyses suggest no puberty-specific effect on audiovisual integration. Gender differences were observed, requiring further study.

Generalized additive mixed models (GAMMs) capture nonlinear developmental patterns and may clarify puberty's role by comparing models with and without pubertal modulation. An age-matching analysis examines individuals at different pubertal stages but the same chronological age to isolate puberty's effects. Sample size constraints may limit statistical power in certain phases and genders.

These findings suggest puberty, alongside chronological age, drives functional maturation in the central auditory pathway, influencing adolescent neural speech processing. Further analyses integrating hormone data and age-matching will refine our understanding of puberty's role in sensory brain development and multisensory integration.

ID: 146

Abstract

Keywords: Acute stress, Stress modulation, two photon microscopy, extracellular electrophysiology, auditory perception

How Acute Stress Modulates Neural Dynamics in the Auditory Cortex?

Lior Dor[1], Ghattas Bisharat[1], Jennifer Resnik[1,2]

[1]Department of Life Sciences, Ben-Gurion University of the Negev; [2]Zelman Center for Brain Science Research, Ben-Gurion University of the Negev;

liordo@post.bgu.ac.il

Stress can exert a significant influence on higher-order cognitive functions, including decision-making, learning, selective attention, and memory. While the impact of acute stress on complex cognitive processes is more commonly studied, its effect on auditory processing and perception remains less explored. To address this gap, we used in vivo two-photon calcium imaging and extracellular electrophysiology in mice, enabling us to monitor changes in neural activity within the primary auditory cortex under both baseline and acute stress conditions. This approach allowed us to characterize how acute stress dynamically modulates sensory processing. Our findings suggest that acute stress leads to a modulation in sound-evoked neural activity in the auditory cortex when compared to baseline, providing new insights into how stress modulates sensory processing and perception.

ID: 147
Abstract
Keywords: Crossmodal Plasticity, Perinatal Deafness

Network Control Theory Reveals Altered Energy Efficiency in the Deaf Brain

Danial Ghiaseddin[1,3], Yaser Merrikhi[1], Alessandra Sacco[2], Stephen G. Gordon[2], Stephen G. Lomber[1,3]
[1]Department of Physiology, Faculty of Medicine and Health Sciences, McGill University, Montreal, Canada; [2]Integrated Program in Neuroscience, McGill University, Montreal, Canada; [3]Department of Psychology, College of Science, The University of Texas at Arlington, Arlington, USA; danial.ghiaseddin@mail.mcgill.ca

Perinatal deafness is associated with enhanced visual abilities, particularly in motion detection and peripheral vision. This phenomenon is attributed to crossmodal plasticity, where auditory brain regions become involved in visual processing. However, diffusion MRI (DTI/DSI) studies indicate that only 0.21% of all connections are unique to either the hearing or deaf group, suggesting that large-scale structural rewiring is minimal. The functional implications of these subtle network changes are unclear. One key aspect of these functional changes is the energy consumption required for information propagation between brain regions, which reflects how efficiently neural signals flow through the brain network. To investigate whether differences in structural connectivity influence the energy efficiency of information propagation in the deaf brain, we applied dynamic system models and network control theory. In this study, 12 domestic cats (6 hearing, 6 perinatally deafened) were examined, with deafness induced through ototoxic aminoglycoside antibiotics within the first four weeks after birth. Hearing status was confirmed using auditory brainstem responses. Imaging data were collected using a 7 Tesla MRI scanner, including diffusion-weighted imaging (DWI) for white matter tract reconstruction and T1-weighted MP2RAGE scans for anatomical reference. Preprocessing steps included brain extraction, distortion correction, and alignment to a standardized feline gray matter atlas (Catlas). Following the extraction of the structural connectome for each subject, the brain was modeled as a dynamical system where regions (nodes) are connected by white-matter tracts (edges), and the minimum control energy required for information to propagate between regions was calculated. This approach quantifies energy consumption using the structural connectome as a constraint. Our analysis suggests that while large-scale structural differences are limited, functional network dynamics in the deaf brain may be subtly reconfigured. Simulations using network control theory revealed information propagation from the visual cortex to the prefrontal cortex required 1.06% less energy in the deaf brain ($p = 0.01$). Likewise, information transfer from the visual cortex to frontal regions required 0.84% less energy ($p = 0.01$), suggesting that higher-order cognitive regions become more efficiently accessible following auditory deprivation. Similarly, visual-to-motor transitions required 0.61% less energy ($p = 0.03$), showing a trend toward increased integration between sensory and motor processing. In contrast, information propagation from the visual to auditory cortex required 0.71% more energy ($p = 0.001$), suggesting that auditory deprivation leads to reduced efficiency in engaging auditory circuits. Notably, these observed energy alterations exceed the 0.21% difference in structural connectivity, reinforcing the impact of even minor structural changes on network efficiency.

ID: 149

Abstract

Keywords: VEP, Cat, Deafness, Auditory Cortex, Visual Cortex

Time Course of Cortical Plasticity Following Adult-Onset Deafness in Cats

Siyu Zhu[1,2], Xiaohan Bao[1], Stephen G. Lomber[1,2]

[1]McGill University, Montreal, QC, Canada; [2]University of Texas at Arlington, Arlington, TX, United States; siyu.zhu2@mail.mcgill.ca

When one sensory modality is lost, the brain frequently adapts and enhances the capabilities of the remaining senses. This compensation can occur through cross-modal plasticity, in which brain regions previously dedicated to the missing sense are rewired to support the intact sensory systems, or through intramodal plasticity, where sensory loss can lead to functional reorganization within the spared sensory modalities. Compensatory plasticity has been widely regarded to stem from early and complete sensory deprivation during the critical periods of brain development, hence plasticity after adult-onset deafness has not been well documented. However, an extracellular recording study on ferrets with adult-onset deafness revealed somatosensory conversion within the auditory cortex as early as 16 days post-deafening (Allman et al., 2009), demonstrating a potential for compensatory plasticity even in maturity. In our investigation, we aimed to further explore plasticity after adult-onset deafness on a macro scale using the well-established cat model and studied the time course of its emergence by characterizing changes in motion-onset visually evoked potentials (VEPs). We hypothesized that compensatory plasticity would occur in cats even after adult-onset deafness, and that this reorganization would be reflected by an amplification of VEPs that would continue to increase over time. To test our hypothesis, four cats were ototoxically deafened in adulthood ($\bar{x}$ = 5.2y) using a combination of kanamycin and furosemide, and deafness was confirmed through a flat click ABR at 80dB SPL both during and at 1 month post deafening. VEPs were acquired before and after the deafening of each animal. Mann-Whitney U Tests, Tau-U tests, and permutation tests were performed to compare VEP waveforms during the deaf time course to the hearing baseline. Over a span of roughly 400 days, we observed a progressive amplification in motion-onset VEPs that plateaued by approximately 300 days, evidenced by significantly ($p < .05$) enhanced signal power and peak amplitude. Additionally, we identified a trend of topographical variation in VEP change between recording site locations, reflected by VEP amplification and a significant ($p < .05$) linear trend of decrease in peak latencies over time. The amplification of VEPs was more rapidly observed in recordings from the active subdermal electrode positioned over occipital cortex, whereas changes in latency were more quickly captured by the electrode placed over temporal cortex. Through VEPs, our study evidenced plasticity after deafness even in adulthood and suggests a difference in location and time course for which these changes develop. Overall, VEPs in adult-deafened cats proved to be a valuable tool for exploring the mechanisms underlying the divergent developmental trajectories in normal and hearing-deprived brains, providing evidence suggesting cross-modal and intramodal plasticity in the mature auditory system.

ID: 150
Abstract
Keywords: fMRI, auditory cortex, speech perception, audiovisual integration

Neural and perceptual benefits of audiovisual speech in second-language learners

Sebastian Puschmann[1], Hanin Karawani[2], Pascale Sandmann[1], Esther Ruigendjik[3], Christiane Thiel[4]
[1]Department of Human Medicine, University of Oldenburg, Germany; [2]Department of Communication Sciences and Disorders, University of Haifa, Israel; [3]Institute of Dutch Studies, University of Oldenburg, Germany; [4]Department of Psychology, University of Oldenburg, Germany; sebastian.puschmann@uol.de

Visual cues, such as a speaker's articulatory movements, facilitate speech perception in challenging listening conditions. In native speakers, visual speech cues modulate neural activity across multiple stages of the auditory sensory processing hierarchy, but its effects in second-language learners remain unclear. We here used fMRI and a naturalistic speech-in-noise listening task to investigate audiovisual interactions at different levels of the auditory sensory processing hierarchy in second-language learners.

Twenty-five second-language learners of German (age: 28 ± 5 years, language proficiency: B1 - C1) listened to six unscripted monologues under different listening conditions (speech without noise, speech in noise, audiovisual speech in noise). After each monologue, they rated speech intelligibility and answered ten multiple-choice comprehension questions. Functional MRI data were preprocessed with fMRIPrep, and mean signal time courses were extracted from eight cytoarchitectonically defined auditory cortex regions. Time courses were pruned from physiological noise using nuisance regression and high-pass filtering, and submitted to inter-subject general linear model (GLM) analyses to investigate modulations of stimulus-driven auditory cortex processing by visual speech cues.

Audiovisual speech-in-noise yielded higher intelligibility ratings ($t(24) = 2.9$, $p(FDR) < 0.01$) and comprehension scores ($t(24) = 2.7$, $p(FDR) < 0.05$) compared to auditory-only speech-in-noise. Audiovisual enhancements of stimulus-driven activity were observed in the planum temporale (Te2.2, $p(FDR)<0.05$), the superior temporal gyrus (Te3, $p(FDR)<0.05$), and the upper bank of the superior temporal sulcus (STS1, $p(FDR)<0.01$). Furthermore, functional connectivity revealed stimulus-driven associations between auditory cortex regions (Te3, STS1) and visual areas (V4, V5) during audiovisual speech-in-noise processing (i.e., versus auditory-only speech in noise, all $p(FDR)<0.05$).

Comparisons with a reference group of native German speakers (N=50, age: 26 ± 4 years) that completed the same experimental procedures revealed no significant differences in perceptual benefits, neural audiovisual gains, or connectivity patterns between groups. Furthermore, audiovisual gains were unrelated to second-language proficiency. This suggests that native speakers and second-language learners benefit similarly from visual speech cues.

ID: 152

Abstract

Keywords: pupillometry, belief updating, auditory oddball, prediction error, neuromodulation

Pupil-linked arousal tracks belief updating in dynamic auditory environment

Lars Kopel[1], Evi van Gastel[1], Peter Murphy[2], Simon van Gaal[1], Mototaka Suzuki[1], Jan Willem de Gee[1]

[1]University of Amsterdam, The Netherlands; [2]Maynooth University, Ireland; l.e.kopel@uva.nl

Most of our decisions are guided by predictions—for instance, opting for a bike over a bus based on expected travel time and weather. However, unexpected changes, such as roadworks causing delays, can invalidate these predictions. According to normative Bayesian theory, highly surprising outcomes should prompt rapid belief updating, incorporating new information to improve future decisions. Influential computational models suggest that the locus coeruleus plays a key role in tracking unexpected changes and facilitating belief updating.

We developed a novel belief updating protocol suitable for human and mouse observers, based on a passive oddball design (standard, 90%; oddball, 10%) using high- and low-frequency tones (2kHz and 1kHz for humans; 20kHz and 10kHz for mice; tone duration, 0.5s; inter-tone-interval, 1s). Critically, we introduced (i) two states, defined by their opposite mapping between the frequencies and probabilities of the two tones, and (ii) a 5% chance ("hazard rate") of a state change after every tone presentation. This ensured that both tones appeared equally often within a block, and that oddball status was dependent on the local state. Throughout, we measured pupil size at constant luminance as a marker of ascending arousal, including noradrenergic activity. We hypothesized that pupil responses to the same tone (e.g., low frequency) would be larger when it functioned as a state-dependent oddball vs standard.

Human participants (N=20) were exposed to the passive auditory belief updating protocol. As predicted, tone-evoked pupil response magnitude was larger for state-dependent oddballs than state-dependent standards. Specifically, there was a significant interaction between the effects of tone identity and state on the tone-evoked pupil response (2-way repeated measures ANOVA $F_{1,19} = 39.9$; $p < 0.001$). With a "many controls experiment" we addressed the alternative hypothesis that this result was solely due to stimulus-specific adaptation. Interestingly, the same protocol in the visual domain (using clockwise- and counter-clock-wise-rotated Gabor patches) did not result in larger tone-evoked pupil responses to state-dependent oddballs.

The observed state-dependent differences in pupil response magnitude to the same auditory stimulus led us to the conclusion that after state changes participants update their belief about which tone to expect next. In ongoing work, we aim to characterize the involvement of specific neuromodulatory systems in belief updating, using multi-fiber photometry in primary auditory areas in mice and brainstem fMRI (7T) in humans. Our results will provide the much-needed empirical data to close the gap between computational theory of belief updating and its underlying biological mechanisms.

ID: 154
Abstract
Keywords: aging, auditory cortex, fMRI, hemispheric specialization, training

Changes in auditory cortex activity in older and younger adults during training of listening tasks with different hemispheric involvement

Jörg Stadler[1], André Brechmann[1], Nicole Angenstein[1,2]
[1]Leibniz Institute for Neurobiology, Combinatorial NeuroImaging Core Facility (CNI), Magdeburg, Germany; [2]Otto von Guericke University Magdeburg, University Clinic of Otolaryngology, Head and Neck Surgery, Germany; nicole.angenstein@med.ovgu.de

The lateralization of processing in the auditory cortex (AC) varies for different acoustic features. Thus, the processing of complex acoustic stimuli such as speech requires the involvement of the AC in both hemispheres and thus efficient interhemispheric interaction. Recently, we have shown that older adults require a stronger involvement of the AC for the same performance than younger adults during auditory processing especially in a task that strongly requires interhemispheric interaction (Stadler et al. 2023). This is probably due to reduced anatomical interhemispheric connections in the elderly. In the present fMRI study, we investigated the effect of training on AC activity during two tasks in a group of older (N=17; 57-73 years) and younger adults (N=18; 18-38 years). One task was the categorization of the direction of pitch change of linearly frequency modulated tones which mainly involves the right AC. The other task was the sequential comparison of pitch direction in the same tones, which requires the involvement of both hemispheres and thus a higher degree of interhemispheric interaction. The stimuli were presented monaurally without and with contralateral white noise. An increase in activity during ipsilateral tone presentation due to contralateral noise indicates the main location of task processing (e.g., Stadler et al. 2023, Brechmann & Angenstein 2019). Each participant performed three fMRI-sessions for each task. The task difficulty was adjusted individually to avoid effects of different performance levels on lateralization. In both groups, the performance improved over sessions, even though no feedback was provided. Diffusion-Tensor-Imaging (DTI) revealed lower fractional anisotropy in the fibre tracts connecting left and right AC in the older compared to the younger adults. Furthermore, the size of the corpus callosum was smaller in the older than in the younger group. Functional connectivity was stronger in older compared to younger adults between the AC and various brain regions involved in e.g., working memory, attention, and language processing, particularly during the comparison task. Connectivity between left and right AC was weaker in the older than in the younger adults, possibly due to the reduced anatomical connectivity in older adults. The involvement of the AC increased or decreased over the three sessions, depending on the task, age group and hemisphere. In the categorization task, the involvement predominately decreased in both age groups. In the comparison task, the changes were more variable. Even in the third session, auditory involvement was still stronger in the older than in the younger participants. Although training alters AC involvement differentially in younger and older adults, the differences between age groups remain, despite similar task performance.

This study was supported by DFG (AN 861/4-2).

ID: 156
Abstract
Keywords: speech-in-noise; stream segregation; development; adolescence

Neural signatures of stream segregation: from childhood to adulthood

Elena Benocci[1], Claude Alain[2], Axelle Calcus[1]
[1]Université libre de Bruxelles, Belgium; [2]Rotman research institute, Baycrest Health Sciences; axelle.calcus@ulb.be

When faced with noisy environments, listeners perform auditory scene analysis, which allows them to selectively focus on the relevant auditory target while ignoring interferences. Stream segregation involves organizing similar sound waves into a coherent stream, while distinguishing dissimilar acoustic components and attributing them to distinct sources. Two event-related potential components have been identified as "neural signatures" of stream segregation: the Object-Related Negativity (ORN) and the P400.

Our study aims to examine (i) the maturation of neural correlates of stream segregation and (ii) the development of the relationship between these neural correlates and speech perception in noise. ORN/P400 were recorded while 8-23 year-olds (n = 75) performed an active stream segregation task based on temporal coherence. Participants also performed speech identification in noise tasks (behaviourally). Behavioral results indicate an improvement in both stream segregation and speech perception in noise from childhood to adulthood. Amplitude of the ORN (but not P400) decreases, and latency of both ORN and P400 decreases throughout development. Critically, P400 amplitude significantly predicts stream segregation performance.

Overall, our results suggest that the neural mechanisms underlying stream segregation follow a prolonged maturation trajectory, and support the progressive maturation of auditory scene analysis and speech perception in noise.

ID: 158

Abstract

Keywords: event segmentation, narrative discourse, prosody, speech comprehension, aging

Event processing during listening: The relative contributions of discourse structure and prosody

Yulia Lamekina[1], Ryan A. Panela[1,2], Björn Herrmann[1,2]

[1]Rotman Research Institute, Baycrest Academy for Research and Education, Canada; [2]Department of Psychology, University of Toronto, Canada; ilamekina@research.baycrest.org

The continuous stream of our daily experience is mentally encoded as discrete events, a process referred to as event segmentation. An event represents a coherent sequence of actions and goals, consisting of a salient beginning and an end. Event segmentation is driven by the inherent narrative structure, but it is unclear whether external acoustic parameters of a spoken speech are a contributing factor. In this study we investigated the impact of discourse structure and prosodic cues (here pause duration) on event segmentation.

In the first part of the study, we analyzed pause durations of existing audiobooks and podcasts to investigate whether speakers naturally produce longer pauses at event boundaries than event centers (used as a control). To this end, we segmented a corpus of 200 spoken narratives using a large language model (LLM). Our results show that speakers produce longer pauses at event boundaries than at event centers. This could either reflect real-time event processing by the speaker or a subliminal trend for the speaker to facilitate the event segmentation for the listener. We further investigated speech generated through artificial voices. While older generations of voices display a more uniform distribution of pause lengths, newer models seem to capture the lengthened pause durations at event boundaries.

In the second part of the study, we investigated the influence of pause duration during real-time event segmentation in younger and older adults. Participants listened to spoken stories under three conditions: the original story (unmanipulated), a story for which the pauses at event boundaries were lengthened, and a story for which the pauses at event centers were lengthened. Our results show that pause lengthening at event boundaries did not improve boundary detection. However, pause lengthening at the event centers increased erroneous segmentation responses. Older adults showed greater event segmentation abilities demonstrated by an increased number of responses. Furthermore, we found that participants tended to use information at the end of an event rather than at the beginning of a new event to segment the story. This possibly speaks against the theoretical premise that prediction errors driven by the context at the onset of a new event guides segmentation.

Overall, our study reveals the interplay of discourse and prosodic information in event segmentation. In the presence of thematic narrative shifts, prosodic information does not appear to further contribute to event segmentation. However, prosody can bias listeners into segmenting a narrative in the absence of these shifts, suggesting that prosodic information is used during narrative event processing. The current approach and results provide a new way of studying naturalistic speech comprehension by focusing on the ability to extract larger meaningful units in speech.

ID: 159
Abstract
Keywords: Cochlear Implant, Listening Effort, Voice Training, Speech Perception, Speech in Noise

The Effect of Voice Training on Speech-on-Speech Intelligibility and Listening Effort in Cochlear Implant Users

Ada Biçer[1,2], Thomas Koelewijn[1,2], Deniz Başkent[1,2]

[1]Department of Otorhinolaryngology/ Head and Neck Surgery, University Medical Center Groningen, University of Groningen, The Netherlands; [2]Graduate School of Medical Sciences, Research School of Behavioural and Cognitive Neurosciences, University of Groningen, The Netherlands; t.koelewijn@rug.nl

Understanding speech in a multi-talker situation can be challenging and effortful especially for cochlear implant (CI) users. In normal-hearing listeners, it was previously shown that listening to familiar or trained voices can improve speech intelligibility and/or listening effort in speech-on-speech listening. Recently, we showed that listening effort was reduced for trained voices compared to untrained voices, but only for vocoder processed speech. The aim of this study was to investigate if such benefits can also be provided to CI users with an explicit voice training of the target speaker presented in a speech masker.

Sixteen CI users, ages ranging between 20 and 78 (Mean age = 62.81, SD = 17.37), performed an explicit voice training via a speaker recognition task. Following voice training, speech-on-speech intelligibility, in percent correct scores, and listening effort, using pupillometry, were measured with the Dutch Child-friendly Coordinate Response Measure (CCRM) test. The CCRM test involved a target sentence that consists of a call sign, a color, and a number (e.g. “Show the dog where the blue five is”), uttered by the trained voice and an untrained voice. The masker speech consisted of combined short randomized fragments of CCRM sentences, presented at 0 dB or +6 dB target-to-masker ratios (TMRs), and with shifted voice cues of fundamental frequency (F0) and vocal-tract-length (VTL) combined, at either 6.29 or 12.59 semitones (st) relative to the target voice.

Results from a generalized linear mixed model analysis showed that, overall, voice training did not lead to an intelligibility improvement. Nevertheless, a 3-way interaction showed that listening to trained voices led to better intelligibility when TMR was increased from 0 dB to +6 dB, and when voice cue differences between the target and masker were increased from 6.29 st to 12.59 st. Pupillometry results showed significantly smaller pupil responses for trained voices than untrained voices, at the 0 dB 6.29 st condition, implying a benefit of voice training by reducing listening effort. However, pupil dilation responses were also smaller when listening to untrained voices than trained voices at +6 dB 6.29 st condition.

Outcomes from this study demonstrated that listening to a newly trained voice might lead to an intelligibility benefit when listening situations are the most optimal (+6 dB 12.59 st) for CI users. Nevertheless, less favorable listening situations might lead to a training benefit by a decrease in listening effort in the absence of an intelligibility benefit. Less favorable listening situations can also lead to better intelligibility for newly encountered voices, perhaps via making a better use of F0+VTL voice cue differences between the target and masker when intensity cues are not available (TMR = 0 dB). Together, these results imply that like normal hearing listeners, CI users might also benefit from listening to trained voices under certain conditions.

ID: 160

Abstract

Keywords: prediciton, cortical speech tracking, encoding models, speech, EEG

Cortical speech tracking is modulated by linguistic predictability

Alexandra K Emmendorfer[1,2], Lars Riecke[1], Lars Hausfeld[1]

[1]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, The Netherlands; [2]Section Teaching and Learning Innovation, Faculty of Psychology and Neuroscience, Maastricht University, The Netherlands; a.emmendorfer@maastrichtuniversity.nl

Neural responses have been shown to encode acoustic and linguistic features of continuous speech, and this tracking is sensitive to cognitive factors, such as attention and comprehension. However, little is known about its sensitivity to predictions from repetition. Here, we experimentally manipulate the linguistic and acoustic predictability of speech signals by repeatedly exposing participants to intelligible or unintelligible versions of the same audiobook. Using EEG, we examine whether increasing the familiarity with the linguistic and acoustic content of the speech signal affects the cortical encoding of the speech. In 12 blocks, participants listen to 5 consecutive repetitions of the same 30-s audiobook segment, presented as either clean or 1-band noise-vocoded speech, while performing an auditory attention task. We modelled the cortical encoding of the speech with a linear model including predictors for the envelope and its derivative, phoneme onsets, and 14 phonetic features (initial model comparisons revealed no benefit of adding lexical features such as word surprisal). To test whether the neural encoding of speech is affected by repetition, we compare model performance on left-out data from the repetition on which the model was fitted vs. data from the other repetitions. For clean speech only, we find a significant difference, suggesting that speech encoding varies across repetitions. To verify that this effect reflects predictions, we test whether model performance changes with the temporal distance between the training and testing data. This was confirmed with a linear mixed effects model with the difference in model performance as the outcome variable, and distance from the training data as a predictor. For clean speech, we observe a significant distance effect, where model performance decreases with increasing temporal distance. We find no such effect for noise-vocoded speech, suggesting that it is specific to speech predictability, rather than general factors related to the repeated exposure. Finally, we examine the time course of this distance effect in the temporal response functions encoding the envelope, its derivative, and phoneme onsets. In comparison to repetition 1, repetitions 3 and 4 showed enhanced model weights at 40–130ms and 70–130ms for the derivative, respectively, while repetition 5 showed increased model weights at -90 to -10ms for the envelope, indicative of predictive processing. In summary, we show that cortical tracking of sublexical speech features is sensitive to the linguistic predictability of clean speech. With increasing exposure to the same input, the neural encoding of these features changes gradually. This seems to be driven by cortical response modulations in an early time window (-90 to 130ms), in line with perceptual enhancement of predicted input. These observations are only present for clean speech, confirming that they are specific to changes in predictability of the speech signal.

ID: 161

Abstract

Keywords: Multiscale Auditory Processing, Convolutional Recurrent Neural Network, Deep Lerning, Environmental Sound Recognition, Brain Alignment

Modeling Auditory Cortical Processing: A Deep Learning Approach to Hierarchical Auditory Integration

Michele Esposito[1,3], Giancarlo Valente[1], Bruno L. Giordano[2], Elia Formisano[1,4]

[1]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, the Netherlands; [2]Institut des Neurosciences de La Timone, CNRS UMR 7289 – Université Aix-Marseille, Marseille, France; [3]Institute of Data Science, Faculty of Science and Engineering, Maastricht University, Maastricht, The Netherlands; [4]Maastricht Centre for Systems Biology and Bioinformatics, Faculty of Science and Engineering, Maastricht University, Maastricht, The Netherlands; m.esposito@maastrichtuniversity.nl

Human audition effortlessly integrates brief transients (e.g., a door knock) and sustained textures (e.g., rain) across multiple timescales, a capacity tied to hierarchical temporal integration in the auditory cortex. To capture these diverse neural sensitivities to acoustic dynamics, we present a novel Multiscale Convolutional Recurrent Neural Network (MsCRNN) architecture that operates in parallel on three temporally distinct “buffer” streams (e.g., 50 ms, 200 ms, 400 ms). Within each stream, time-distributed convolutions extract localized spectro-temporal features, an attention layer focuses on salient frequency or temporal cues, and stacked Gated Recurrent Units (GRUs) track evolving patterns over time. The outputs from these parallel streams are then merged at 10 ms resolution, enabling fine-grained and continuous classification of environmental sounds.

We show, through extensive grid-search optimizations, that this multiscale design significantly improves framewise recognition compared to single-scale baselines across a wide range of natural stimuli. Notably, shorter buffers excel at detecting impulsive events like “snaps” or “ticks,” while longer buffers capture slower modulations in sustained sounds such as “wind” or “music.” In addition, the inclusion of attention mechanisms within each scale-sensitive stream substantially enhances the model’s ability to amplify relevant acoustic features and suppress background noise.

Beyond improving classification, the MsCRNN’s time-resolved outputs naturally align with electrophysiological signals (e.g., intracranial EEG or MEG), enabling the possibility for representational similarity or encoding/decoding analyses. By comparing the model’s latent representations to neural data in millisecond increments, researchers can probe how populations of neurons simultaneously encode transient onsets and protracted spectral contours.

Overall, our results offer two key contributions. First, they support the hypothesis that parallel, scale-specific processing streams reflect a neurobiologically inspired strategy for efficient sound analysis. Second, they provide a computational framework for linking machine audition with neuroscience by directly comparing model representations to ongoing cortical dynamics. Such synergy holds promise for advancing both automatic sound recognition and our understanding of the multiscale structure of auditory perception in the human brain.

ID: 162
Abstract
Keywords: Plasticity, Elevation Perception, Earmolds, Spectral Cue Relearning, Localization

Adaptation rate and persistence across multiple sets of spectral cues for sound localisation

Paul Friedrich
Leipzig University, Germany; paul.friedrich@uni-leipzig.de

The adult auditory system adapts to changes in spectral cues for sound localization. This plasticity was demonstrated by modifying the shape of the pinnae with molds. Previous studies investigating this adaptation process have focused on the effects of learning one additional set of spectral cues. However, adaptation to multiple pinna shapes could reveal limitations in the auditory systems' ability to encode discrete spectral-to-spatial mappings without interference and thus help to determine the mechanism underlying spectral cue relearning.

In the present study, 15 listeners learned to localize sounds with two different sets of earmolds within consecutive five-day adaptation periods. To establish both representations in quick succession, participants underwent daily sessions of gamified sensory-motor training. Acoustic and behavioral effects of the earmolds were recorded, as well as the trajectories of individual adaptation to the modified pinnae throughout the experiment. To test whether learning a new set of spectral cues interferes with a recently learned mapping, the persistence of the initial adaptation was measured after participants adapted to the second set of earmolds. Earmolds were removed after each adaptation period, and participants' localization accuracy was immediately measured again to test for aftereffects of the adaptation.

Both pinna modifications severely disrupted vertical sound localization, and participants recovered within each 5-day adaptation period. After the second adaptation listeners were able to access 3 different sets of spectral cues for sound localization. Learning a second set of modified cues did not interfere with the previous adaptation and adaptation rate did not increase with repeated cue relearning. Participants' localization accuracy with their native ears remained unchanged once the molds were removed.

Modified pinna shapes were sufficiently different to cause repeated disruption of vertical localization. Learning rates and adaptation persistence exceeded those observed in previous studies. Adaptation persistence did not differ between the successive earmolds, suggesting that adaptation to the second mold did not interfere with the previously learned representation. Participants adapted to both sets of spectral cues with equal success, indicating a pre-attentive process which does not underlie metaplasticity or procedural learning. The ability to store multiple sets of spectral cues without interference suggests a surprisingly large capacity for this representation, which may be indicative of a cortical participation.

ID: 163
Abstract
Keywords: perceptual learning, speech perception, cortical tracking

Neuronal Mechanisms Supporting the Perceptual Learning of Degraded Speech

Ellie Ambridge[1], Edmund Lalor[2], Ediz Sohoglu[1]
[1]University of Sussex, United Kingdom; [2]University of Rochester, USA;
E.Sohoglu@sussex.ac.uk

Human speech perception has a remarkable capacity to cope with sub-optimal auditory stimuli. This ability to cope depends, in part, on perceptual learning i.e. relatively long-lasting improvement in understanding degraded speech due to past experience or training. Two divergent theories of perceptual learning have been proposed. Transformation mechanisms suggest that listeners learn to reverse the effect of degradation by a process of compensation or inverse transformation. This contrasts with a cue-reweighting mechanism which suggests a reweighting of acoustic-phonetic cues, for example, upweighting of intact cues and downweighing of degraded cues.

In the present study, 30 normal hearing listeners (25 female; mean age 21.47) were trained over three days to understand spoken sentences in which fine spectral modulations or fast temporal modulations were filtered out. Participants also listened to a clear speech control condition whereby the acoustic signal was unchanged. From day one to three, participants' word report accuracy increased from ≈ 20% to ≈ 60%, showing robust perceptual learning. On days one and three, we used EEG and temporal response function analysis to assess neural tracking of intact modulations i.e. those present in both clear and filtered speech. We also assessed neural tracking of degraded speech modulations i.e. those present in clear speech but attenuated in filtered speech.

We found that neural tracking of both intact and degraded modulations increased in day three versus day one. This effect was specific to filtered speech with no change observable for the clear speech control condition.

These results provide evidence in favour of a transformation-based account of perceptual learning. Our findings shed important insights into how the brain adapts to perceptually challenging stimuli, with possible future clinical implications for cochlear implant users. In a follow-up experiment, we are investigating whether the present findings generalise to more naturalistic listening situations i.e. during perception of audiobook speech.

ID: 164

Abstract

Keywords: human auditory cortex, fMRI, spectrotemporal processing, speech, vowels

Using vowels to dissociate the representations of three distinct spectrotemporal features within and beyond human auditory cortex

Yongtian Ou[1], Kendrick Kay[1,2], Andrew Oxenham[1]

[1]Department of Psychology, University of Minnesota, United States of America; [2]Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota, United States of America; ou000036@umn.edu

Spectral content, fundamental frequency (F0), and spectrotemporal modulation are three key acoustic features of speech. They are known to be encoded in auditory cortex, but less is known about the relative strength and spatial distribution of their representations, or how they are represented beyond auditory cortex. To investigate representations of the three acoustic features, we designed a stimulus set based on spoken vowels, which vary with respect to all three features and may help drive brain responses due to their ecological relevance... We measured ultra-high-field (7T) fMRI responses to single spoken vowels that systematically varied with respect to vowel identity (probing spectral content), speaker identity (probing F0), and pitch glide direction (probing spectrotemporal modulation). Three regions showed selectivity to these features: auditory cortex, premotor regions, and the inferior parietal sulcus (IPS). Voxel-wise selectivity maps revealed that the strongest selectivity to vowel identity was spatially focused in primary auditory regions, whereas speaker identity and pitch glide direction showed weaker, more dispersed selectivity in non-primary regions. In contrast, premotor regions and IPS displayed comparable selectivity across all three features. Multivoxel pattern analysis (MVPA) demonstrated that all three regions could successfully decode the three features, with auditory cortex showing the highest accuracy, particularly for vowel identity. Unlike auditory cortex, premotor and IPS regions did not exhibit a prominently higher accuracy for vowel identity than the other two features. These findings suggest that spectrotemporal features of spoken vowels are represented both within and beyond auditory cortex. While representations in auditory cortex are strongest for spectral content over F0 and spectrotemporal modulation, non-auditory regions showed similar selectivity for spectral content and the higher-level features of F0 and spectrotemporal modulation.
[Supported by NIH grant R01 DC005216.]

ID: 165

Abstract

Keywords: computational modelling, connectivity, event-related fields, evolutionary algorithms, human auditory cortex

Estimating connectivity patterns in the human auditory cortex via computational modelling

Ewelina Tomana[1], Nina Härtwich[2], Mateusz Lickindorf[1], Reinhard König[2], Patrick J. C. May[3,2], Cezary Sielużycki[1]

[1]Wrocław University of Science and Technology, Wrocław, Poland; [2]Leibniz Institute for Neurobiology, Magdeburg, Germany; [3]Lancaster University, Lancaster, United Kingdom; ewelina.tomana@pwr.edu.pl

Investigating the complex connectivity and signal propagation within the auditory cortex (AC) of healthy humans poses a significant challenge due to the restriction to noninvasive techniques such as magnetoencephalography (MEG). MEG offers valuable insights into auditory processing, but the interpretation of the acquired signal is difficult because the measured magnetic field is a global and indirect measure of neural activity in the AC. Additionally, factors such as low signal-to-noise ratio and inter-subject variability complicate the analysis.

We hypothesise that AC connectivity is, at least to some extent, reflected in the morphology of auditory event-related fields (ERFs) and that a computational model of the AC can be used to extract connectivity information from these ERFs. Our approach builds upon the modelling strategy by May and colleagues: 1) the AC model simulates MEG signals based on synaptic currents, 2) the core, belt, and parabelt areas, as well as the auditory fields therein, are explicitly represented, and 3) signal propagation is governed by feedforward and feedback connections.

To generate synthetic ERFs that closely match experimental data, we iteratively adjust connection strengths. Our optimisation procedure is based on an evolutionary algorithm and we focus on MEG data capturing the adaptation phenomenon in the human AC. The resulting model connectivity patterns are predictions of in-vivo anatomy, with commonalities found across subjects hinting at universally applicable aspects of AC organisation. Analysis across a population revealed that feedback connections between core and belt regions dominate over feedforward connections in both hemispheres, with a stronger asymmetry observed in the left hemisphere. Preliminary results from an alternative optimisation method—sequential neural posterior estimation—support these findings.

Additionally, we used the optimised AC model to examine the generation and adaptation of the separate ERF components arising from activity in the core, belt, and parabelt areas. All of these areas contributed to ERF generation, indicating that the ERF does not originate from spatially separated sources. Instead, it reflects the activity of the entire AC. Moreover, we found that the contribution from each area adapts, suggesting that adaptation does not emerge due to a reduction of activity in separate sources but rather due to a more global change in AC dynamics.

Our study highlights the critical role of parameter estimation in computational modelling, demonstrating how refined models can uncover previously hidden information in experimental data. Future work will explore alternative optimisation strategies to ensure robustness and reliability, paving the way for deeper insights into AC connectivity and its functional implications.

ID: 166

Abstract

Keywords: spatial hearing, signal discrimination, d-prime, JND, psychometric function

Rethinking JNDs: Evidence Against Signal Detection Theory in Sound Lateralisation

Jakab Pilaszanovich[1], Kirke Müller[1], Fulvio Domini[2], Marc Schönwiesner[1]

[1]Leipzig University, Germany; [2]Brown University, USA;

jakab.pilaszanovich@gmail.com

Just-noticeable differences (JNDs) are commonly assumed to depend on the variability of two signals being compared. In the current study, we show that this may not be the case. In a two-alternative forced choice (2AFC) task participants judged whether a second sound was perceived to the left or right of the first. Reference and comparison stimuli were lateralised using either interaural time difference (ITD) or interaural level difference (ILD) cues, and compared across all possible pairings. Contrary to standard predictions, JNDs were not influenced by the variability of the reference stimulus. Instead, participants' sensitivity depended only on the comparison cue. This result challenges a core assumption in probabilistic models of perception and suggests that widely used methods for calculating sensory discrimination may need to be reconsidered.

ID: 167
Abstract
Keywords: Language, Attention, Predictive, EEG, Model

Neural tracking of phonetic predictability under varying attentional load

Dominic West, Jeff Mitchell, Sophie Forster, Ediz Sohoglu
University of Sussex, United Kingdom; dsw27@sussex.ac.uk

Predictive processing is a theoretical framework that proposes the brain continuously generates predictions about incoming sensory input. It suggests that perception is not a passive process of receiving stimuli but an active process of minimising prediction error – the difference between expected and actual sensory input. Previous neuroimaging studies have reached opposing conclusions as to whether attentional focus is required for the predictive processing of speech. One explanation for these conflicting results is that the tasks used in experiments have varied in their level of perceptual difficulty or 'load'. According to the Perceptual Load Theory of attention, in a task with low perceptual load, spare brain resources can 'spill over' to process task-irrelevant stimuli. In a task with high perceptual load, no spare capacity is available and task-irrelevant stimuli remain unprocessed.

In this study we manipulate perceptual load using a visual task whilst participants listen to an audiobook. The task requires participants to track either two ('low load') or five ('high load') moving dots within a larger set of identical dots. Trials last 10 seconds and the audiobook plays continuously throughout each ~8 minute block. We also include a control condition whereby participants are instructed to attend to the audiobook and ignore the visual task. In this condition, participants are instructed to press a key whenever they detect an occasional repeated sentence in the audiobook. We use EEG to measure changes in the neural tracking of speech predictability between conditions. In this way, we can directly assess the relationship between attention and predictive processing in the context of speech comprehension.

To assess neural tracking of speech features we regress changes in phoneme surprisal against the EEG response to the audiobook and use 'leave-one-trial-out' cross-validation to compute the accuracy of each model. Phoneme surprisal represents how surprising a phoneme is, given the words that are consistent with the unfolding speech signal and their relative frequency in the language. We also compute acoustic models derived from the broadband envelope and the rectified derivative of the broadband envelope. Individual models demonstrate reliable neural tracking, as inferred by comparing observed model accuracies with those obtained after random shuffling of the EEG data. Unexpectedly, in preliminary analyses (N=30), we find no reduction in the neural tracking of phoneme surprisal when participants attended to visual compared with auditory input, nor when comparing low and high perceptual load during the visual task. In ongoing analyses, we will investigate whether speech predictive processing is modulated by attention in higher linguistic levels (e.g. word surprisal).

ID: 168

Abstract

Keywords: auditory steady-state response, syllabic, phonemic, neural synchronization

Measuring neural synchronization to syllabic and phonemic speech envelopes using nested modulations

Ana Belen Carbajal Chavez[1], Maaike Smit[1], Marlies Gillis[1], Maaike Vandermosten[1], Pol Ghesquière[2], Jan Wouters[1]

[1]Experimental Oto-rhino-laryngology, Department of Neuroscience, KU Leuven, Belgium; [2]Parenting and Special Education, KU Leuven, Belgium; ana.carbajalchavez@kuleuven.be

The strength at which the brain phase-locks to the speech envelope provides insight into speech processing. This neural synchronization can be measured through neurophysiological recordings of the auditory steady-state response (ASSR), elicited by low-, single-frequency amplitude-modulated (AM) sounds. However, due to high interindividual variability in neural synchronization —especially below 20 Hz— single-frequency modulations provide limited insight into phase-locking across an entire frequency band. For example, assessing neural synchronization at 4 Hz does not fully capture synchronization dynamics within the theta band, which corresponds to the syllabic rate in natural speech.

Recently, Gransier and Wouters (2021) developed the Temporal Envelope Speech Tracking (TEMPEST) framework, which utilizes speech-like stimuli to assess a distribution of modulation frequencies rather than a single discrete frequency. Previous findings demonstrated a strong correspondence between neural activity evoked by TEMPEST stimuli and that elicited by AM sounds at syllabic (~4 Hz) and phonemic (~20 Hz) rates, suggesting that TEMPEST provides a more comprehensive measure of speech temporal processing compared to single-frequency ASSRs (David et al., 2022). Leveraging on the TEMPEST framework, we developed nested-TEMPEST, a novel stimulus that integrates nested distributions of phonemic and syllabic modulation frequencies, enabling simultaneous assessment of both frequency bands allowing the measurement time to be significantly reduced.

This study aims to evaluate the reliability and robustness of nested-TEMPEST by (1) assessing intra-subject correspondence of neural activity elicited by AM sounds, single distribution TEMPEST and nested-TEMPEST stimuli, (2) evaluating the test-retest reliability of the nested-TEMPEST across two sessions, and (3) investigating to which extend neural activity elicited by nested-TEMPEST is related to neural tracking to continuous speech in silence (N = 25). Our goal is to validate the nested-TEMPEST as a time-efficient, reliable, and standardized framework for measuring speech processing. Data collection is ongoing, and first results will be presented at the conference.

References

David, W., Gransier, R., & Wouters, J. (2022). Evaluation of phase-locking to parameterized speech envelopes. Frontiers in Neurology, 13, 852030. https://doi.org/10.3389/fneur.2022.852030

Gransier, R., & Wouters, J. (2021). Neural auditory processing of parameterized speech envelopes. Hearing Research, 412, 108374. https://doi.org/10.1016/j.heares.2021.108374

ID: 169

Abstract

Keywords: behavioral oscillations; beta; speech; lexical tone; entrainment;

Behavioral beta oscillations are visible for the categorization of Mandarin Chinese words manipulated in lexical tone and frequency

Thu Ngan Dang[1,2], Andrea E. Martin[1,3], Sanne Ten Oever[1,3,4]

[1]Max Planck Institute of Psycholinguistics, Nijmegen, The Netherlands; [2]University of Cambridge, United Kingdom; [3]Donders Centre for Cognitive Neuroscience, Nijmegen, The Netherlands; [4]Maastricht University, The Netherlands; sanne.tenoever@maastrichtuniversity.nl

Behavioural oscillations have been observed during the categorization of ambiguous Dutch words in which perception is rhythmically biased to the more or less statistically frequent response option. Here, we aimed to demonstrate similar behavioural oscillations during the categorization of ambiguous Mandarin words manipulated in lexical tone and lexical frequency. We conducted three behavioural experiments in which participants listened to an ambiguous morph after the presentation of an entrainment chain at 6.25 Hz or 10 Hz, or after a single noise burst. We found significant perceptual modulation around low beta frequency range (13 – 19 Hz) for the noise burst experiment, around the entrainment frequency for the 10 Hz experiment, and an absence of behavioural oscillations for the 6.25 Hz experiment. Interestingly, all effects happened for the word pair with bigger lexical frequency differences, but in which the direction of the lexical tone and lexical word frequency difference did not align. The results are consistent with models of weakly coupled oscillators assuming a natural frequency around the low beta range for the lexical categorization task. They also suggest that lexical tone frequency was not of importance for our reported effect. Our study contributes to the findings that neuronal oscillations bias the categorization of ambiguous stimuli that vary in statistical frequency and - considering the outcome at the beta range – could point to a role of the motor system in lexical tone categorization.

ID: 170

Abstract

Keywords: change deafness, auditory attention, attentive filters, corollary discharge, subjective perception, self-generation

The fundamental frequencies of our own voice

Hakam Neamaalkassis[1], Yves Boubenec[2], Christian J. Fiebach[1], R. Muralikrishnan[3], Alessandro Tavano[1]

[1]Goethe University Frankfurt, Frankfurt a.M., Germany; [2]École Normale Supérieure, PSL Research University, Paris, France; [3]Max Planck Institute for Empirical Aesthetics, Frankfurt a.M., Germany; neamaalkassis@psych.uni-frankfurt.de

Own actions send a corollary discharge (CD) signal, that is a copy of the planned motor programme, to sensory-specific brain areas to suppress the anticipated sensory response, providing a neural basis for the sense of self. [1] When we speak, the sensory consequences of the fundamental frequency ($f0$) of our own voice, generated by vocal fold vibrations, are suppressed [2,3]. This effect is known as speaking-induced suppression (SIS), a special case of motor-induced suppression (MIS). However, due to bone/air conduction filtering effects, the $f0$ we self-generate is measurably different from the $f0$ we subjectively perceive as defining our own voice.[4,5] Accordingly, we hypothesised 1) a reduced frequency discrimination resolution in the vicinity of one's $f0$, as a long-term byproduct of frequent SIS in response to self-generated speech; and 2) similar effect magnitudes for the subjectively and objectively measured $f0$, assuming that the transfer function between both $f0$s is known to the brain. Using an auditory local odd-ball paradigm with multiple deviation magnitudes, we parametrically tested the sensitivity to auditory change in the frequency neighbourhoods of objective and subjective own voice pitch $f0$s, as well as a control $f0$. A mixed-effects model of the behavioural data showed that participants experience change deafness for both own-pitch $f0$s to a similar extent, relative to a control pitch condition. Additionally, neither the estimated marginal means nor pairwise contrasts between the subjective and objective $f0$ conditions produced a significant difference in the behavioural profiles. We conclude that when we listen attentively, we are likely to filter out small pitch changes in the vicinity of our own objective and subjective voice $f0$, possibly as a long-term consequence of speaking-induced suppression mechanisms integrated with individual, perceptual bodily priors.

References:

[1] Crapse TB, Sommer MA. 2008 Corollary discharge across the animal kingdom. Nat. Rev. Neurosci. 9, 587–600. (doi:10.1038/nrn2457)

[2] Eliades SJ, Wang X. 2008 Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature 453, 1102–1106. (doi:10.

1038/nature06910)

[3] Eliades SJ, Wang X. 2003 Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. J. Neurophysiol. 89, 2194–2207. (doi:10.1152/jn.00627.2002)

[4] Kimura M, Yotsumoto Y. 2018 Auditory traits of 'own voice'PLoS One 13, e0199443. (doi:10.1371/journal.pone.0199443)

[5] Reinfeldt S, Östli P, Håkansson B, Stenfelt S. 2010 Hearing one's own voice during phoneme vocalization—transmission by air and bone conduction. J. Acoust. Soc. Am. 128, 751–762. (doi:10.1121/1.3458855)

ID: 171
Abstract
Keywords: non-primary auditory cortex, prefrontal cortex, neuroanatomy, ferret

Tracking brain-wide projections to and from the ferret's non-primary auditory cortex

Zheyi Tang, Rebecca Norris, Clem Hallou, Stephen Town, Jennifer Bizley
University College London, United Kingdom; zheyi.tang.23@ucl.ac.uk

Our ability to focus on specific sounds while filtering out background noise, known as auditory selective attention, is essential for distinguishing relevant acoustic signals in complex environments. This process requires multiple brain regions beyond the auditory cortex, including the prefrontal, motor, and parietal cortices. The ferret is a popular animal model for auditory neuroscience research, and neurophysiological recordings indicate a hierarchy of areas in which attention and behavioural context shapes auditory cortical responses to look increasingly like those observed in dorsolateral frontal cortex. These areas include secondary tonotopic and non-tonotopic fields within the Posterior Ectosylvian Gyrus (PEG). However, the anatomical connections between the secondary areas on the PEG and non-auditory regions, particularly the prefrontal cortex, remain poorly defined.

To better understand the neural circuits that may facilitate attentional modulation of auditory cortical activity, we performed tracer experiments using viral vectors and cholera toxin subunit B (CTB) labelling. Retrograde viral tracers expressing fluorescent reporters were injected into the PEG (n=2 ferrets), while CTB was injected into prefrontal subregions (n=5 ferrets, one with additional anterograde virus) to assess reciprocal connectivity. Neurons and axons were imaged using widefield and confocal microscopy, with labelled neurons and terminal fields mapped to the ferret brain atlas (Radtke-Schuller, 2018).

Retrograde labelling revealed substantial inputs to the PEG from multiple brain regions: In the prefrontal cortex, labelled cells were found in the dorsal prefrontal cortex, orbital gyrus, and medial prefrontal cortex, including the dorsal anterior cingulate cortex and prelimbic cortex. Additional inputs originated from the primary and premotor cortices, hippocampus, and multisensory areas such as the medial bank of the rostral suprasylvian sulcus and the anterior and posterior medial-lateral suprasylvian cortex. Anterograde tracing revealed projections from AC to all auditory cortical fields, the auditory thalamus and inferior colliculus. Prefrontal CTB injections revealed only very sparse labelling in the auditory cortex. However, combined analysis with PEG injection data identified several brain regions with robust connectivity to both prefrontal areas and PEG, including the orbital gyrus, premotor cortex, parietal cortex, and cingulate cortex. Notably, the full rostro-caudal extent of the cingulate cortex exhibited reciprocal connections with prefrontal cortex and projected to the PEG.

Ongoing work is expanding this dataset utilising varied tracing approaches and more restricted viral injections to more completely understand the non-auditory inputs to secondary auditory cortex.

ID: 172

Abstract

Keywords: Two-Photon Microscopy, Mouse, Population Coding, Learning, Stimulus generalization

Representational maps in the auditory cortex

Takahiro Noda[1], Dominik F Aschauer[1], Jens-Bastian Eppler[2], Matthias Kaschube[2], Yonatan Loewenstein[3], Simon Rumpel[1]

[1]University Medical Center, Johannes Gutenberg University Mainz, Germany; [2]FIAS, Goethe University Frankfurt, Germany; [3]The Hebrew University, Jerusalem, Israel; sirumpel@uni-mainz.de

The advent of large-scale recordings of stimulus-evoked activity patterns in rodent models has allowed the application of concepts and analytical techniques originally developed in human cognitive studies using fMRI. One such method, Representational Similarity Analysis (RSA), is a powerful tool for assessing the pairwise similarities of population responses, enabling the estimation of the structure of a representational map within neuronal activity space. Here, we applied RSA to chronic two-photon calcium imaging data recorded from awake, passively listening mice, capturing responses from tens of thousands of neurons to a range of pure tones and complex sound stimuli. When using one of the complex sounds as a conditional stimulus in an auditory cued fear conditioning paradigm, we found that the representational map allowed to predict specific learning-induced changes for non-conditioned sounds in the stimulus set located nearby [1]: At the level of neuronal activities, we found a higher likelihood of sounds driving the same neuronal ensemble as the conditioned sound, suggesting a form of neuronal association. Behaviorally, this was reflected by a high degree of stimulus generalization. Additionally, targeted laser microablation of few functionally identified neurons temporarily impaired the representational map, but this impairment gradually recovered over several days [2]. By analyzing the underlying changes in single neuron tuning, we discovered an active, homeostatic recruitment of neurons that were previously unresponsive to the sound set. Together, our findings demonstrate that neurons in the auditory cortex collectively arrange the activity patterns evoked by various sounds in an active manner such that the similarities between different activity patterns reflect the perceptual relatedness of the corresponding stimuli.

[1] Aschauer et al. (2022) Learning-induced biases in the ongoing dynamics of sensory representations predict stimulus generalization. Cell Reports, 38:110340

[2] Noda et al. (2025) Homeostasis of a representational map in the neocortex. Nat. Neurosci., accepted

ID: 173

Abstract

Keywords: expectation, prediction, loudness, physiology, aversiveness

Conditioned and social predictions of sound intensity and related aversiveness

Ester Benzaquén[1], Dheerendra, Pradeep[1,3], Timothy D. Griffiths[1], Sukhbinder Kumar[1,2]

[1]Newcastle University, United Kingdom; [2]The University of Iowa, Iowa; [3]University of Glasgow, United Kingdom

ester.benzaquen@newcastle.ac.uk

Expectations can modify how we perceive and sense the world. Predictions could be formed by our previous experience or associative learning, or by social information relayed by others. Here, we induce predictions of loudness and associated aversiveness, and record behavioural, physiological, and neural measures of expectation and perception. Using visual cues, the loudness (in dBs SPL - High: 90; Medium: 86; or Low: 82) of a forthcoming 1 kHz tone could be predicted. A conditioned CS+ cue was followed by the High or Medium tone half of the time, while a CS- was followed by the Low or Medium tone half of the time. A social cue, signifying the supposed/fabricated aversive ratings of 10 previous participants to the forthcoming tone, was also presented. This social cue could predict an overall high or low aversive tone, but it did not bear any relation to the loudness of the upcoming tone. Data from 44 participants demonstrated that the visual cues induced subjective expectancy as measured by expectancy ratings of loudness. Furthermore, while only analysing medium loudness tones, the predicted intensity of the sound was associated with the perceived aversiveness, such that tones that were predicted as louder were rated as more unpleasant even though loudness was the same. Measures of arousal and autonomic function further showed that this effect transferred to participant's physiology: sounds that were predicted as louder induced a faster galvanic skin response and greater heart deceleration. The same paradigm was tested in 4 participants awaiting recession surgery for the treatment of epilepsy. Intracranial recordings further demonstrated that predictions of aversiveness can modify the neural processing of sounds in the auditory cortex. Taken together, our data shows that predictions of loudness can shift the perceived aversiveness of sounds in line to expectations and modify their physiological and neural processing.

ID: 174

Abstract

Keywords: Cochlear Implants, Speech Perception, Objective Measures, Auditory Processing

Phase-locked neural responses to temporal modulations are associated with speech perception in CI users

Elise Verwaerde[1], Wouter David[1], Julian Schott[1,2], Robin Gransier[1], Jan Wouters[1]

[1]ExpORL, Dept. of Neurosciences, KU Leuven, Herestraat 49/721, 3000 Leuven, Belgium; [2]STADIUS, Dept. of Electrical Engineering, KU Leuven, Kasteelpark Arenberg 10/2440, 3001 Leuven, Belgium; elise.verwaerde@kuleuven.be

Temporal envelope modulations (TEMs) are essential cues for speech perception in cochlear implant (CI) users, who rely heavily on temporal information due to limited spectral resolution. Understanding how the auditory system encodes these modulations is key to optimizing speech outcomes. Electrically evoked auditory steady-state responses (eASSRs) offer an objective, non-invasive measure of neural sensitivity to amplitude-modulated (AM) stimuli, capturing phase-locked responses across (sub)cortical auditory regions.

This study investigated the potential of eASSRs to assess neural encoding of TEMs across the electrode array in adult CI users and to guide individualized programming strategies. We recorded EEG to 40-Hz eASSRs evoked by 900 pulses per second (pps) AM stimuli presented in monopolar mode on six intracochlear electrodes. Data were collected across three sessions using a BioSemi hyper-rate sampling system (262 kHz sampling rate), enabling more accurate modeling of stimulation artifacts. Growth functions were derived for each electrode to determine objective stimulation thresholds across the dynamic range.

Based on these neural markers, six experimental stimulation programs (MAPs) were designed per subject. Threshold and comfort levels of the MAPs were set using eASSR-derived and behavioral measures. Speech perception was evaluated for each MAP and compared to each participant's clinical program using standard speech tests in quiet and in noise.

Results showed that good-performing CI users achieved comparable speech understanding with the experimental MAPs and their clinical program. However, a modest performance advantage was observed with the familiar clinical settings. These findings suggest that eASSRs can provide meaningful, site-specific insight into neural temporal processing and support individualized fitting approaches. Ongoing research in CI users with poorer performance aims to identify more pronounced variability in neural encoding profiles and their relation to speech outcomes.

This work supports the use of eASSRs as a tool to investigate temporal processing along the electrically stimulated auditory pathway and their potential clinical application in objective CI programming. More broadly, it contributes to our understanding of cortical and subcortical processing of temporal cues in implanted listeners.

Acknowledgements:

This work was partly funded by Cochlear Technology Centre Belgium and the Flanders Innovation & Entrepreneurship Agency through the VLAIO research grant [HBC.2019.2373; HBC.2020.2308], and partly by an SB Ph.D. grant from the Research Foundation Flanders (FWO) awarded to EV [1SF5123N].

ID: 175
Abstract
Keywords: Emotional Prosody, Speech perception, EEG

Characterizing the impact of emotional prosody on the neural processing of continuous speech

Asena Akkaya[1], Martin Winchester[1], Sonia Bishop[1,2], Giovanni Di Liberto[2]
[1]Trinity College Dublin, Ireland; [2]UC Berkeley; akkayaa@tcd.ie

Prosody plays a crucial role in conveying the meaning of speech (Paulmann et al., 2011). Identifying emotions through vocal cues is a fundamental component of social cognition and human communication. In this study, we investigated how two key dimensions (emotional prosody and social relevance) shape the neural encoding of continuous speech. Emotional prosody refers to the modulation of pitch, rhythm, and intensity that conveys emotional states, while social relevance pertains to whether speech occurs in a dialogue (socially interactive) or a monologue (solitary speech).

This study takes the first step in that direction when considering continuous speech listening scenarios. The neurophysiology of continuous speech listening has been extensively investigated in recent years. However, no previous studies have isolated the impact of emotional prosody on the neural tracking of continuous speech. Here, we investigate that relationship by presenting participants (N=20) with continuous speech stimuli (dialogues and monologues) vocalized with various emotional tones, including happiness, sadness, fear, and anger. Neural activity was recorded non-invasively using electroencephalography (EEG), along with skin conductance, heart rate, and behavioral measurements of emotional prosody detection. Multivariate lagged regression analysis was employed to examine the impact of emotional prosody on the neural processing of speech. The resulting temporal response functions reveal a separation of the emotional and social components across the EEG bands delta and theta. Specifically, emotional prosody affected the neural encoding of speech in the EEG delta-band, while social relevance impacted the speech encoding in the theta-band. Notably, anger elicited the strongest delta-band effect. These findings provide novel insights into how emotional prosody and social relevance influence the neural tracking of continuous speech, advancing our understanding of speech perception in naturalistic settings.

ID: 176
Abstract
Keywords: Speech Production, Naturalistic, TRF, EEG

Investigating neural predictive mechanisms during spontaneous speech production

Martin M. Winchester, Giovanni M. Di Liberto
Trinity College Dublin, Ireland; winchesm@tcd.ie

Speech production relies on the integration of linguistic, cognitive, and motor processes, yet the neural mechanisms underlying lexical prediction remain insufficiently understood. This study investigates the role of predictive mechanisms in lexical selection. We designed an experiment to probe lexical processing during speech production, testing the hypothesis that predictive processing supports lexical selection, where the selection of more unlikely words would correspond to a higher computational cost, resulting in larger neural activations. Crucially, we test this hypothesis in scenarios involving the production of continuous speech segments, offering insights into how the brain generates predictions prior to articulation in naturalistic settings.Participants performed two tasks while EEG and EMG were recorded: a constrained spontaneous speech task (taboo) and an automatic speech repetition task. These two tasks were selected to disentangle neural activity reflecting spontaneous lexical selection versus the generation of memorised speech. In the taboo task, participants described target words while avoiding related "taboo" words, a condition designed to elicit high surprisal and cognitive effort. In the repetition task, participants listened to and repeated pre-generated descriptions, where lexical surprisal and entropy were matched to a typical taboo session. EEG data was analysed using temporal response functions (TRFs) to model neural responses to acoustic and lexical features.Preliminary findings indicate that lexical predictions are encoded in pre-onset neural activity, with distinct components observed during the constrained lexical selection task compared to repetition. This suggests that these metrics reflect the cognitive demands of lexical retrieval and may serve as markers for mental effort in speech production. The results also highlight the utility of EEG combined with TRF analysis in capturing fine-grained neural dynamics during naturalistic tasks, despite the methodological challenges posed by articulatory artifacts.This study bridges the gap between predictive processing in speech perception and production, contributing to a growing body of evidence that statistical expectations shape neural activity in auditory cognition.

ID: 177

Abstract

Keywords: Sound offset, speech-in-noise, temporal processing, sound onset, auditory processing

Sound Offset Encoding is Related to Speech-in-Noise Perception at Sentence Level in Older Adults

Hasan Colak[1,2], Xiaoxuan Guo[1], Adèle Simon[3], Ester Benzaquén[1], William Sedley[4], Jennifer F. Linden[3,5], Timothy D. Griffiths[1,6,7]

[1]Biosciences Institute, Newcastle University, Newcastle upon Tyne, NE2 4HH, United Kingdom; [2]Department of Audiology, Hacettepe University, Ankara, Türkiye; [3]Ear Institute, University College London, London, United Kingdom; [4]Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University; [5]Department of Neuroscience, Physiology, & Pharmacology, University College London, London, United Kingdom; [6]Department of Imaging Neuroscience, University College London, London, United Kingdom; [7]Human Brain Research Laboratory, University of Iowa, Iowa City, IA, USA.; h.colak2@newcastle.ac.uk

Sound onset and offset encoding are critical aspects of auditory temporal processing with important roles in speech-in-noise perception. We investigated brain correlates of active-offset encoding in older adults and their relationship to real-world listening. 34 adults aged 60–81 years underwent pure-tone audiometry, and assessment of speech in noise ability at both word and sentence levels. EEG data were recorded while participants performed active duration discrimination of noise segments and during passive listening. Onset and offset responses were extracted from EEG data, and amplitude differences were analysed. We found robust onset and offset responses at a single-subject level. No significant difference was observed in sound-onset responses between active and passive listening conditions. Significantly larger sound-offset amplitudes were demonstrated in the active compared to the passive condition. The finding supports an attentional mechanism for sound-offset encoding. We found that active sound-offset amplitudes were significantly correlated with sentence-in-noise performance, while no correlation was observed for word in noise. We suggest a role for offset analysis in parsing of the speech stream needed for speech segregation. The active EEG measure we use allows single-subject inference and is a potential clinical measure of auditory cognition relevant to speech-in-noise listening.

ID: 178

Abstract

Keywords: audiovisual, virtual-reality, modeling, neural-tracking, electroencephalography

Exploring Objective Measures of Audiovisual Benefit in Speech-in-Noise Using Modeled Inferior Colliculus Responses

John Cooper, Jonas Vanthornhout, Astrid van Wieringen, Tom Francart

ExpORL, Department of Neurosciences, KU Leuven, Belgium;
johnkyle.cooper@kuleuven.be

When listening in a noisy environment, seeing the mouth of a speaker can significantly improve a listener's speech comprehension (Sumby and Pollack, 1954). Numerous research studies have investigated the neural pathways that contribute to this multisensory processing (Peelle and Sommers, 2015). One of these pathways involves the inferior colliculus (IC), a subcortical structure in the auditory system. There is evidence from previous research that there is retinal innervation in the IC (Gruters and Groh, 2012).

Additionally, previous research has developed models that simulate the responses from the IC to various sounds (Nelson and Carney, 2004). The responses from these models have been shown to improve the prediction of electroencephalography (EEG) responses to speech signals (Lindboom et al, 2023). The prediction of EEG responses from a feature of a speech signal has been coined as neural tracking, which has been shown to correlate with speech intelligibility (Vanthornhout et al., 2018). Objective predictions of speech intelligibility for individual participants has high value as this paves the way toward clinical applications of neural tracking that would improve hearing outcomes for individuals with hearing loss.

To our knowledge, there is no research that investigates if these IC model responses can predict audiovisual benefit from seeing the mouth movements of a speaker while they are speaking. Therefore, we investigate this using a recently recorded dataset from Cooper et al, 2025 that recruited 20 Dutch-speaking participants to listen to semantically valuable sentences in varying signal to noise ratios (SNRs) ranging from -9 dB SNR to silence and with or without the presence of a virtual avatar speaking these sentences. Using this dataset we investigate if the neural tracking of the IC model responses to the sentences is a good predictor of audiovisual benefit for speech in noise. To accomplish this, we measure neural tracking by using linear regression to estimate the predictive ability of the IC model responses.

Based on the results from Cooper et al, 2025 which shows that the neural tracking of the speech envelope may predict behavioral audiovisual benefit on a group level, we expect that the neural tracking of the IC model responses will yield similar results. Moreover, we hypothesize that the increased predictive capabilities of the IC model responses will improve the relationship between the behavioral and neural tracking measures of audiovisual benefit for the individual participants.

ID: 179
Abstract
Keywords: auditory cortex

Sound offset responses become highly informative in the auditory cortex.

Charly Lamothe, Sophie Bagur, Etienne Gosselin, Brice Bathellier
Pasteur Institute, France; brice.bathellier@cnrs.fr

The entire auditory system downstream of the cochlea features pronounced offset responses which follow the termination of sounds. Due to their ubiquity, it remains an unsolved question whether offset responses are generated early in the auditory system and subsequently propagated downstream or recomputed at each processing stage. Here, we analysed large-scale sound response datasets acquired in the cochlear nucleus, inferior colliculus, medial geniculate nucleus and auditory cortex of awake mice. All brain régions showed a significant proportion of offset responses, often combined with onset and sustained responses in the same neuron. However, using population activity decoders, we observed that neural representations after the sound offset show a three-fold increase in sound encoding accuracy in the cortex relative to subcortical areas. This result indicates that cortical offsets encode a more precise short-term memory of the elapsed sound than subcortical offsets and that they likely result from specific computational steps.

ID: 180
Abstract
Keywords: audition, cortex, mouse, population coding, sleep

The geometry of cortical sound processing in slow wave sleep

Allan Muller, Sophie Bagur, Brice Bathellier
Institut Pasteur, Institut de l'Audition, France; allan.muller@pasteur.fr

Sensory disconnection in sleep remains an elusive phenomenon. It was proposed, based on observations under anesthesia, that it could derive from a disrupted geometry of sound responses which merge with the spontaneous activity subspace. Here, we followed large neural populations of the mouse auditory cortex across slow wave sleep and wakefulness. We observed that sleep dampens sound responses but preserves the geometry of sound representations which remain separate from spontaneous activity. Moreover, response dampening was strongly coordinated across neurons and varied throughout sleep, spanning from fully preserved response patterns to population response failures on a fraction of sound presentations. These failures are more common during high spindle-band activity and are not observed in wakefulness. Therefore, in sleep, the auditory system preserves sound feature selectivity up to the cortex for detailed acoustic surveillance, but concurrently implements a gating mechanism leading to epochs of reduced responsiveness and even local sensory disconnections.

ID: 181

Abstract

Keywords: Speech comprehension, Competing speech, Temporal response functions, EEG

Neural Signatures of Attentional Engagement Predict Speech Comprehension in Multitalker Environments

Alice Vivien Barchet[1,2,3], Andrea Bruera[1], Jasmin Wend[1], Johanna M. Rimmele[4], Jonas Obleser[5,6], Gesa Hartwigsen[1,2]

[1]Research Group Cognition and Plasticity, Max Planck Institute for Human Cognitive and Brain Sciences; [2]Cognitive and Biological Psychology, Leipzig University; [3]International Max Planck Research School on Cognitive Neuroimaging; [4]Department of Cognitive Neuropsychology, Max Planck Institute for Empirical Aesthetics; [5]Department of Psychology, University of Lübeck; [6]Center for Brain, Behavior, and Metabolism, University of Lübeck; barchet@cbs.mpg.de

In everyday life, spoken speech streams are often masked by noise or competing speech streams in the surroundings, creating challenging listening situations. In these situations, listeners adaptively use selective attention to filter the sensory input. Previous studies have demonstrated that selective attention shapes the neural representation of lower- and higher-level speech features. Attended streams are typically represented more strongly than unattended ones, suggesting either enhancement of the attended or suppression of the unattended stream. However, it is unclear how these complementary processes support attentional filtering and speech comprehension on different hierarchical levels.

In this study, we used multivariate temporal response functions to analyze the EEG signals of 43 young adults, examining the relationship between the neural tracking of acoustic and higher-level linguistic features in target and distractor streams and a fine-grained speech comprehension measure. The participants listened to 240 dual-talker trials, each consisting of two sentences presented simultaneously. Participants attended to one speaker while ignoring the other and repeated the target sentence. The individual task difficulty was adjusted using an adaptive staircase procedure to account for peripheral hearing differences. We assessed comprehension on a word-by-word level, predicting comprehension accuracy using word-level stimulus variables and neural tracking measures.

Behaviorally, comprehension was enhanced for words with higher audibility ($z = 23.19$, $p < .001$) and lower surprisal ($z = -18.24$, $p < .001$). Additionally, there was a significant interaction between audibility and surprisal ($z = 9.66$, $p < .001$) suggesting that more predictable words were easier to understand than less predictable words when it was harder to hear the attended stream. Brain-behavior correlations showed that the neural tracking of word and phoneme onsets ($z = 4.03$, $p < .001$) and word-level linguistic features ($z = 2.22$, $p = .026$) in the attended stream predicted comprehension at the individual single-trial level. Moreover, acoustic tracking of the ignored speech stream was positively correlated with comprehension performance ($z = 2.20$, $p = .027$), whereas word level linguistic neural tracking of the ignored stream was negatively correlated with comprehension ($z = -2.63$, $p = .008$).

Collectively, our results suggest that lower-level and higher-level mechanisms jointly support comprehension performance in multitalker environments. Attentional filtering during speech comprehension requires target enhancement as well as distractor suppression at different hierarchical levels.

ID: 182
Abstract
Keywords: frequency-modulated tones, learning, magnetic resonance imaging, memory, rodent

Effects of punishment on performance and auditory-evoked brain activation patterns in discrimination learning – an auditory fMRI study in the Mongolian gerbil

Annika Michalek[1,2], Patricia Wenk[1], Nicole Angenstein[1,3], Eike Budinger[1,4]
[1]Leibniz Institute for Neurobiology, Magdeburg, Germany; [2]Istituto Italiano di Tecnologia, Rovereto, Italy; [3]Otto von Guericke University, Magdeburg, Germany; [4]Center for Behavioural Brain Sciences, Magdeburg, Germany; budinger@lin-magdeburg.de

Discrimination of two auditory stimuli like between a rising and falling frequency-modulated tone(FMT) is a standard behavioral task performed on rodents to disentangle brain processing during learning. In Mongolian gerbils, this task has given interesting insight into the role of auditory cortex in auditory learning. Aversive reinforcement for misses and false alarms (FA) is often applied to enhance discrimination learning; however, for some animals it causes difficulties in performing the task. For better understanding of reinforcement effects, we performed discrimination training in Mongolian gerbils for FMT direction with and without additional punishment (aP) for FA and parallel fMRI.

30 adult male gerbils were trained to discriminate between rising (GO: 1-2 kHz) and falling (noGo: 2-1 kHz) FMTs in ten daily sessions, enforced through electrical foot shocks for misses; 14 of the animals received aP for FA. Prior, after 3 and 10 sessions, auditory-stimulated BOLD fMRI was performed on the anaesthetized gerbils at 9.4 T with a GE-EPI sequence. For auditory stimulation, GO and noGo tone as well as a control tone were presented in a pseudorandomized block design. The gerbils were divided into “good” and “bad” learners (GL and BL) according to their average learning score (hits - FA rates). Group GLMs were performed on the preprocessed data, statistical comparisons were made on extracted BETA values.

While the learning score did not vary in training with and without aP, hit and FA rates were significantly lower with aP in every session, resulting in significantly more enforcement of the animals and decrease in average learning speed. Activation patterns of GL were similar throughout training with additional BOLD activation in thalamus, retrosplenial and cingulate cortex after 3 sessions; the return to initial activation levels after 10 sessions indicates a consolidation of the task. BL without aP showed a similar activation increase after 3 sessions, but even further growth after 10 sessions. BL with aP did not show as much activation after 3 sessions; after 10 sessions they reached similar activation levels to BL without aP. With aP, a shift of the HRF from tone onset to onset of the miss shock revealed much higher correlation between individual performance and BETA values in isocortex, hippocampal formation, basal ganglia and thalamus after 10 sessions.

Here, we report for the first time the effect of aP in standard auditory discrimination learning on brain-wide activation patterns in rodents. By means of a newly established auditory fMRI we were able to record significant brain activation changes in gerbils caused by FMT discrimination training. Not only did we evidence differences in activation patterns in dependence of their performance during the training, but also distinct variations in patterns due to aP, with more distinguished correlation of individual learning score to brain activation in areas associated with learning and memory.

ID: 183
Abstract
Keywords: Auditory processing, Natural sounds, acoustic-to-semantic, AI-based modeling, RSA

Temporal dynamics of natural sounds representation in the human brain

Marie Plegat[1,2], Giorgio Marinato[1], Christian Ferreyra[1,3], Maria Araújo Vitória[4], Michele Esposito[4], Daniele Schon[2], Elia Formisano[4], Bruno Giordano[1]
[1]Institut de Neurosciences de la Timone, CNRS, Aix-Marseille Univ., Marseille, France; [2]Institut de Neurosciences des Systèmes, Inserm, Aix-Marseille Univ., Marseille, France; [3]Laboratoire d'Informatique et des Systèmes (LIS) CNRS, Centrale Méditerranée, Marseille, France; [4]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht Univ., Maastricht, The Netherlands; marie.plegat@univ-amu.fr

Transforming natural sounds into knowledge of the auditory environment is key for real-world behavior. Although this process involves converting acoustic signals into semantic representations, these aspects are seldom modelled together. Importantly, extensive model comparison approaches have yet to be applied to time-resolved methods like magnetoencephalography (MEG; e.g., De Lucia et al., 2010; Lowe et al., 2023), leaving the computational nature of dynamic cerebral representations unclear. Furthermore, despite recent studies showing how convolutional neural networks (CNNs) best predict perceptual and fMRI representations (e.g., Giordano et al., 2023), the potential of these models at accounting for dynamic cerebral responses to natural sounds remains unexplored.

Here, we addressed these gaps by predicting MEG responses to natural sounds using three classes of computational models: acoustic, text-based semantics, and sound-processing CNNs. In particular, we considered two sound-processing CNNs that differ only in their training objective (Esposito et al., 2024): CatDNN, trained to learn sound-event categories, and SemDNN, trained to learn continuous semantic embeddings (Word2Vec).

We found that CNNs predict brain responses to natural sounds more accurately than both acoustic and text-based semantic models. While acoustic and CNN models show peak performance around 250ms from sound onset, CNNs present stronger predictive power, indicating they capture more than just low-level features for early latencies. Focusing on acoustics, whereas early MEG predictions are driven by Cochleagram and MTF (low-level finely-grained acoustic representations), an Auditory Dimensions model, capturing dimensions of auditory sensation (pitch, loudness, brightness, periodicity and roughness) dominated later latencies (800-1000ms), suggesting higher-level processes, and post-primary anatomical locations (e.g., STG). A detailed analysis of CNNs revealed that the early layers of both networks similarly predict MEG responses at 250 ms from sound onset, with a small advantage for CatDNN. In contrast, later SemDNN layers drove prediction between 500-800ms, suggesting a potential switch from early categorical semantic (/acoustic) representations (early layers of both networks) to later continuous semantic representations. Finally, text-based models, including Word2Vec, captured less accurately MEG responses than CNNs that learn Word2Vec directly from sound (SemDNN), suggesting that natural sound semantics in the brain overlaps only in part with that of linguistic materials.

Overall, our findings provide a temporally resolved computational view of the cerebral representation of natural sounds, and reveal how CNNs can uncover basic structural aspects of natural sound semantics in the human brain.

ID: 184
Abstract
Keywords: sequence, neurophysiology, frontal, auditory, order

Neural encoding of identity and ordinal information during auditory sequences in the macaque frontotemporal cortex

Steven P. Errington[1], Beshoy Agayby[1], Colline Poirier[1], Yaneri A. Ayala[2], Joel I. Berger[2], Timothy D. Griffiths[1,2,3], Christopher I. Petkov[1,2], Yukiko Kikuchi[1]
[1]Biosciences Institute, Newcastle University, Newcastle-upon-Tyne, United Kingdom; [2]Department of Neurosurgery, University of Iowa, Iowa City, United States; [3]Wellcome Centre for Human Neuroimaging, University College London, UK; steven.errington@newcastle.ac.uk

Extracting ordered relationships between auditory events in a sequence is fundamental to communication and cognition. Our previous work demonstrated distinct neural processes in the human and monkey auditory cortex that are sensitive to probabilistic sequencing relationships after the learning of temporal relationships in sequences of speech sounds drawn from an artificial grammar paradigm (Kikuchi et al, 2017, PLoS Biology). Here, we build on this earlier work by investigating more directly how neurons encode both the identity and order of elements within learned auditory sequences. To do this, two adult rhesus macaques performed a passive listening task in which they maintained central fixation while hearing structured sequences of nonsense words. Sequences initially followed rule-based (grammatical) structures and were then followed by sequences that either conformed to or violated the learned grammar. During this task, we recorded single-unit activity using multi-channel linear electrode arrays simultaneously targeting auditory cortex (core fields R and RT, belt field RM) and frontal cortex (areas 44, 45, and the frontal operculum). Across 86 recording sessions, we isolated 1,534 neurons, comprising 851 from the auditory cortex and 683 from the frontal cortex. Across the population of single neurons, principal component analysis (PCA) showed that auditory neuronal trajectories were similar across the individual sequence elements, starting and ending at the same point for each element, suggestive of stable, element-specific encoding. In contrast, trajectories of frontal cortex neurons evolved progressively across the sequence, with each element represented by a distinct PCA space. Similar results were also observed at the level of individual neurons. We identified neurons that were significantly modulated by individual elements within the sequences, revealing diverse response dynamics across auditory and frontal cortices. Clustering of these response patterns revealed four main temporal profiles—facilitated, ramping, suppressed, and peri-onset—that were shared across brain regions but varied in prevalence. Finally, while auditory cortex neurons exhibited clear periodic spiking aligned to the rhythm of sequence presentation, frontal cortex neurons showed more variable, non-periodic activity, suggesting distinct encoding strategies for tracking structured auditory input. Collectively, these findings suggest neurons in the auditory cortex appear to encode information at the element-level within a sequence. In contrast, neurons in the frontal cortex show activity patterns that reflect the position of each element within the overall sequence—indicating a role in tracking order and maintaining structured relationships between elements. These results shed new light on the neural underpinnings of sequence prediction in the primate frontotemporal system.

Support: MRC (MR/X003701/1) & joint NSF (SBE2342847) and BBRSC (BB/Z00005X/1) grants.

ID: 185
Abstract
Keywords: auditory cortex, learning, behavior, categorization, layer 5

Deep-layer projection neurons develop representations of perceptual categories and behavioral choice

Nathan A. Schneider[1,2,3], Michael Malina[2,3,4], Ross S. Williamson[1,2,3,5,6]
[1]Department of Otolaryngology, University of Pittsburgh; [2]Center for the Neural Basis of Cognition; [3]Pittsburgh Hearing Research Center; [4]Neuroscience Institute, Carnegie Mellon University; [5]Department of Neurobiology, University of Pittsburgh; [6]Department of Bioengineering, University of Pittsburgh; nas290@pitt.edu

Auditory-guided behavior is a fundamental aspect of our daily lives, whenever auditory information guides our decisions and actions. A ubiquitous example of this is auditory categorization: the transformation of acoustic stimuli into discrete perceptual categories. While we have previously shown that auditory cortex (ACtx) is necessary for such categorical behavior, the underlying circuit mechanisms remain unknown. Nestled amongst several populations of cells, extratelencephalic (ET) neurons reside primarily in layer (L)5 of ACtx and broadcast auditory information to diverse sub-cortical targets associated with decision-making, action, and reward. To investigate the behavioral role of L5 ET neurons, we trained head-fixed water-restricted mice to categorize the rate of sinusoidal amplitude-modulated (sAM) noise bursts as either fast or slow to receive a water reward. We then used two-photon calcium imaging alongside selective GCaMP8s expression to monitor the activity of L5 ET, as well as L2/3 and L5 intratelencephalic (IT) populations.

In expert mice, L5 ET neurons displayed preferences for slow or fast sAM categories during behavior. This categorical selectivity was not present in untrained mice, and longitudinal recording revealed that L5 ET neurons dynamically shifted their responses across training to reflect these learned categories. Critically, this categorical selectivity was stronger during active behavioral engagement as compared to passive presentation of identical stimuli on the same day, suggesting top-down modulatory input. The strength of the L5 ET categorical selectivity was also correlated with behavioral performance. In contrast, neither L2/3 nor L5 IT neurons displayed categorical selectivity. To ensure these effects truly reflected changes in sensory processing, and not pre-motor contributions, we utilized a generalized linear framework to determine the extent to which individual neurons were influenced by both sound and movement. Removing movement-driven neurons from our analyses did not change our categorical selectivity findings, suggesting that this effect is independent of motor-related activity.

We then investigated how these neural populations encoded behavioral choices. Our task contained a stimulus at the category boundary, which was neither slow nor fast, but rather the rewarded action was evenly split between left and right across trials. This allowed us to investigate choice independent of stimulus categories. Decoding accuracies and PCA trajectories of population activity revealed that all three populations contained information about left/right choice, but this signal was strongest in the L5 ET population. Together, these results suggest that the ACtx L5 ET projection system selectively propagates behaviorally relevant signals to downstream areas and is critical for auditory-guided behavior.

ID: 186
Abstract
Keywords: vocalisations, auditory stimulation, free-moving behaviour, emotional regulation, neuropsychiatric disorders

Modulation of behavioural states in free-moving marmosets using emotionally salient acoustic stimuli

Emma Woolgar[1], Steven Errington[1], Benjamin Slater[1], Christopher Morris[2], Quoc Vuong[1], Timothy G Constandinou[3], Andrew Jackson[1], Yukiko Kikuchi[1]
[1]Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK; [2]Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK; [3]Department of Electrical & Electronic Engineering, Imperial College London, UK; e.woolgar2@newcastle.ac.uk

Neuropsychiatric disorders such as anxiety and autism spectrum disorder are characterised by disruptions in emotional regulation. The amygdala plays a pivotal role in emotional processing, and its interactions with other brain regions underpin a range of emotional and cognitive control processes. Dysregulation of amygdala connectivity, reflected through atypical oscillatory activity, is increasingly recognised as a core feature of these disorders. However, despite the amygdala's importance in emotional regulation, relatively few studies have investigated whether its oscillatory dynamics can be modulated through non-invasive means such as acoustic stimulation. Here, we aim to develop a translational marmoset model to assess whether emotionally salient acoustic stimuli can modulate behavioural readouts associated with amygdala function. Common marmosets (Callithrix jacchus) are highly social non-human primates with a rich vocal repertoire and robust emotional responses to specific auditory cues. Prior studies have demonstrated that marmosets reliably produce 'tsik' alarm calls in response to fear-evoking stimuli and that playback of familiar conspecific tsik calls can reduce behavioural and hormonal stress responses (i.e. cortisol level) in socially isolated individuals (Cross and Rogers, 2004), suggesting a potential calming effect. To further explore the behavioural impact of different call types, we exposed isolated marmosets (N = 3) to three types of conspecific vocalisation: 'phee', 'tsik' and 'twitter'. Animals were placed in a sound-attenuated chamber and underwent four 15-minute sessions of acoustic stimulation or silence (control), with conditions randomised to reduce environmental habituation. Behavioural markers including locomotion and exploratory behaviour and stress indicators such as self-scratching were identified and quantified through BORIS and DeepLabCut. Preliminary results show that acoustic stimulation consistently reduced stationary behaviour and increased activity relative to the silent condition, regardless of vocalisation type. However, exploration and locomotion were more frequently observed during playback of 'phee' and 'twitter' calls compared to 'tsik' calls. Given that exploratory behaviour and movement in novel environments are associated with reduced stress and adaptive emotional regulation, these findings suggest that social calls ('phee', 'twitter') may have a more positive effect than alarm-related calls ('tsik'). This study provides an early proof-of-concept for using emotionally salient vocalisations to influence behavioural markers linked to emotional function. By demonstrating that specific acoustic stimuli with prominent theta-band (4-10 Hz) acoustic features can modulate stress-related behaviours, this work has potential implications for non-invasive approaches, such as combining them with closed-loop acoustic stimulation to study and treat emotional dysregulation in neuropsychiatric disorders.

ID: 187
Abstract
Keywords: Homotopic Coupling, Functional Connectivity, fMRI, Naturalistic Stimulation, Resting State

Naturalistic Stimulation Evokes Homotopic Functional Connectivity Patterns in the Auditory Cortices That Are Not Present in Resting State

Nima Zargarnezhad, Ingrid S Johnsrude
Western University, Canada; nzargarn@uwo.ca

Introduction and Objective

The two brain hemispheres are, to a first approximation, mirror images of each other. Homotopic connectivity is the structural connection between homologous regions of neocortex through the corpus callosum and other commissures. Direct structural connections imply coordinated activity between the (specialized) hemispheres, and perhaps strong functional coupling between homologous regions in the two hemispheres (homotopic functional connectivity HFC) which may be modulated depending on cognitive state or neurological conditions (Mancuso et al. 2019).

In this study, we aim to characterize the functional connectivity between homologous sensory cortices using neuroimaging data, and compare this between resting state (RS) and the more naturalistic movie-watching (MW) conditions.

Materials & Methods

The Human Connectome Project (Van Essen et al. 2013) includes 4 sessions of 15 min of movie-watching (MW) and 4 sessions of 15 min of resting-state (RS) data collected with 7T scanners (TR = 1s) from 79 healthy, unrelated, and young adults (F=41). These were pre-processed with the HCP pipeline (flat-map normalization to the bilaterally symmetric FSLR template; Glasser et al. 2013) and parcellated into 360 Glasser (Glasser et al. 2016) regions-of-interest (ROIs).

For each participant, timeseries from homologous ROI pairs (i.e., the same region in the left and right) were correlated to provide the HFC for that region, for individual sessions in both MW and RS. To examine dynamic HFC (dHFC; Savva et al. 2019), we calculated correlations within non-overlapping sliding windows (length 30 s) for each MW data timeseries. To examine synchrony across participants in this dynamic coupling, the average pairwise intersubject correlation (Hasson et al. 2004) of these dHFC “time series” (2nd order ISC) was estimated for each ROI pair for each MW session.

Results

We observed an increased HFC (>0.15) in MW (>0.65) compared to RS (0.50-0.80) across the auditory cortices. Visual cortices’ HFC was very high in the RS and did not significantly increase with the naturalistic stimulation in MW (both >0.75).

We also observed common patterns of dHFC in the auditory and early visual cortices across participants in the MW condition, suggesting a dynamic and stimulus driven coupling strength in these regions. This observation was supported by relatively high 2nd order ISC values in auditory and early visual regions (>0.25). These results were consistent across the 4 sessions of MW.

Conclusions

In summary, in this study we showed that MW evokes common homotopic coupling patterns across healthy participants in the auditory and early visual cortices that are not present in RS. These common patterns promise a robust, quick, and non-invasive metric to identify abnormal functionality of the auditory cortex in neurological conditions.

ID: 188
Abstract
Keywords: high gamma, human, iEEG, semantic categorization, speech

Transformation of Stimulus-driven into Behavior-driving Activity along the Cortical Auditory Pathways

Kirill V. Nourski[1], Mitchell Steinschneider[1,2], Ariane E. Rhone[1], Matthew A. Howard[1]
[1]The University of Iowa, Iowa City, IA, United States of America; [2]Albert Einstein College of Medicine, Bronx, NY, United States of America; kirill-nourski@uiowa.edu

Auditory areas on the superior temporal plane and lateral convexity are key initial stages of speech processing in the human cortex, representing acoustic and phonetic attributes in a temporally precise manner. More complex representations in auditory-related cortex along the ventral and dorsal processing streams and prefrontal cortex are associated with perception and action. In this study, we used intracranial electroencephalography (iEEG) to clarify where and how stimulus-driven neural activity is progressively transformed into that leading to perceptually driven behavioral events.

Participants were neurosurgical patients undergoing iEEG monitoring for medically intractable epilepsy. Stimuli were monosyllabic words, with participants pressing a button in response to a target category (Nourski et al., J Neurosci. 2022;42:5034-46). High gamma (70-150 Hz) event-related band power (ERBP), normalized to a prestimulus baseline, was measured in two windows: (1) after stimulus onset and (2) before motor response. Three activity patterns were identified: (1) “stimulus-related” (significant ERBP only in the 1st window); (2) “behavior-related” (significant ERBP only in the 2nd window); (3) “intermediate” (significant ERBP in both windows).

More than 6000 recording sites were examined. The stimulus-related pattern was the most common response type throughout the cortical auditory hierarchy. By contrast, behavior-related activity was sparsely represented throughout the brain. It had the highest prevalence in the prefrontal cortex, a more limited representation in anterior temporal and sensorimotor cortex and was practically absent in auditory cortex. Interestingly, the stimulus-related pattern was more common than the behavior-related pattern in sensorimotor cortex. Hemispheric asymmetries included a higher prevalence of behavior-related responses in the left hemisphere and a higher prevalence of stimulus-related activity in the right hemisphere. The behavior-related activity was particularly prominent in pars triangularis of the inferior frontal gyrus, middle frontal gyrus, anterior temporal lobe in the left hemisphere, and in the orbital gyrus bilaterally.

Results reveal progressive stages of cortical auditory processing wherein sensory stimulus-driven responses are ultimately transformed into activity time-locked to subsequent behavior. Future work will test the hypothesis that stimulus-related activity in sensorimotor cortex and behavior-related activity in the left hemisphere are specific for speech.

ID: 189

Abstract

Keywords: noise exposure, parvalbumin neuron, neuroinflammation, auditory cortex

Differential activation of TNFR1 and TNFR2 in cortical parvalbumin-positive neurons following noise trauma

Büşra Köse-Özkan[1,2], Weihua Wang[1], Shaowen Bao[1]

[1]Department of Physiology, College of Medicine, University of Arizona, Tucson, Arizona 85724, USA; [2]Koç University Research Center for Translational Medicine, Istanbul, 34450, Turkiye; bkose19@ku.edu.tr

Noise-induced auditory processing deficits have been linked to alterations in neuroinflammatory signaling within the central auditory system. A key neurobiological mechanism underlying these deficits involves changes in inhibitory neural circuits, particularly those mediated by parvalbumin-positive (PV+) neurons in the primary auditory cortex (AI). However, the role of tumor necrosis factor-alpha (TNF-α) signaling, particularly through its receptors TNFR1 and TNFR2 in PV+ neurons following acoustic trauma remain unclear. Here, we employed a Cre-loxP-based genetic tool to selectively knockdown TNFR1 or TNFR2 in PV+ neurons of the AI in PV-Cre-tdTomato transgenic mice using Cre-dependent viral vectors. Following the transfection period after viral vector injection, mice were exposed to unilateral noise to induce hearing deficits. Behavioral assessments, including the prepulse inhibition and gap detection tests and whole-cell voltage-clamp recordings were conducted before and after noise trauma to evaluate auditory processing and electrophysiological properties of AI pyramidal neurons, respectively. Additionally, immunofluorescence stainings were performed to evaluate parvalbumin expression, microglial density and deramification. Auditory brainstem response thresholds confirmed noise-induced hearing impairment. Immunofluorescence analyses revealed differential effects of TNFR1 and TNFR2 knockdown on PV+ neuron survival and microglial activation. Specifically, TNFR1 knockdown was associated with the preservation of PV+ neurons, suggesting a pro-inflammatory role of TNFR1 in noise-induced damage. Conversely, TNFR2 knockdown exacerbated the increase of microglial density and noise-induced microglial deramification in the AI, highlighting a protective function for TNFR2 signaling. The behavioral findings were consistent with the cellular changes, as TNFR1 knockdown mice did not exhibit noise-induced deficits in gap detection performance observed in control animals. Electrophysiological recordings showed that noise exposure resulted in reduced inhibitory and increased excitatory synaptic transmission, causing an inhibitory-to-excitatory synaptic imbalance. Noise exposure significantly reduced the frequency of miniature inhibitory synaptic currents (mIPSCs) in control and TNFR2 knockdown groups, but not in TNFR1 knockdown mice after noise trauma. These findings support the hypothesis that TNFR1 and TNFR2 differentially regulate PV+ neuron survival and microglial activation in the AI following acoustic trauma and provide novel insight into the molecular mechanisms underlying noise-induced auditory dysfunction.

ID: 190
Abstract
Keywords: humans, EEG, virtual reality, sensorimotor mismatch

MULTIMODAL MISMATCH RESPONSES: a common phenomenon in mouse and human auditory cortex

Magdalena Solyga[1], Georg Keller[1,2]
[1]Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland; [2]Faculty of Science, University of Basel, Basel, Switzerland.; magdalena.solyga@fmi.ch

A major challenge in systems neuroscience is translating insights from animal models to humans, especially in the study of predictive processing. While animal studies have uncovered detailed circuit mechanisms, based on which predictive processing has become a well-established framework for understanding cortical function, direct translation to human research has been limited—partly due to differences in experimental paradigms.

In prior work, we used virtual reality and two-photon imaging to show that audiomotor mismatch responses in mouse auditory cortex are enhanced by concurrent visuomotor mismatches, suggesting that multimodal, non-hierarchical interactions shape prediction error signals in cortical layer 2/3. We now extend this approach to humans.

We developed a virtual reality paradigm for freely moving participants, combining EEG recordings with a closed-loop virtual reality system in which participants can explore a virtual environment by walking around. For audiomotor coupling, we used continuous white noise which volume scaled with walking speed. First, we tested how movement modulates auditory responses - a phenomenon well documented in animals but less explored in humans. Our preliminary data showed enhanced auditory responses during movement compared to stationary conditions, different from findings in mice. To study prediction error responses, we also briefly interrupted the sensorimotor coupling in this environment at random times. We found strong responses to both audiomotor and visuomotor mismatches, that occurred faster than standard sensory-evoked potentials. This suggests that the computational mechanisms of prediction error computations may be shared between mouse and human cortex.

To our knowledge, this is the first demonstration of sensorimotor mismatch responses in humans using walking as a predictive signal. Since predictive processing is often disrupted in individuals with psychiatric or neurological disorders, our research could help develop new biomarkers of disrupted brain function.

ID: 191
Abstract
Keywords: Mice; Auditory cortex inactivation; inferior colliculus; Muscimol; noise; auditory discrimination task

The consequences of inactivating auditory cortex on behavioral discrimination performance in situations of acoustic degradations.

Alexandra Martin, Elise Motsch, Huetz Chloé, Edeline Jean-marc
NeuroPSI, France; alexandra.martin2@universite-paris-saclay.fr

In many species, discriminating vocalizations is an essential component of acoustic communication. Yet animals and humans are constantly exposed to noisy environments. The auditory system is supposed to provide the neural basis for this discrimination, and the literature tends to attribute this ability to auditory cortex (ACx). Indeed, over the last decade, a large number of studies have looked for correlations between auditory cortex responses and behavioral performances in various species and different training paradigms (Narayan et al 2007; Shetake et al 2011; Town et al 2018). This led to the popular concept of robust, invariant, cortical representation of stimulus identity. However, very few studies have tried to compare the quality of neural discrimination obtained when recording cortical versus subcortical neurons.

We evaluated the neuronal discrimination performance of guinea pig cortical neurons and three subcortical structures - the auditory thalamus (MGv), the inferior colliculus (IC) and the cochlear nucleus - at presentation of four conspecific vocalizations presented in quiet and in various levels of noise (Souffi et al 2020, 2021, 2023). Recently, we replicated this experiment in CBA/J mice by recording cortical neurons, MGv and IC neurons at the presentation of those same four guinea pig vocalizations in the same noise levels. Both in guinea pig and mice, the IC neurons showed higher discrimination abilities than the thalamic and cortical neurons, and their responses were more robust to noise than the cortical and thalamic ones.

To go further, we assessed whether behavioral performance indeed relies on cortical or subcortical activity by pharmacologically inhibiting auditory cortex during training. In several mice, discrimination performances were intact both in quiet and in noise during cortical silencing. Together, these data raise questions about the role of auditory cortex during behavioral discrimination performance in situations of acoustic degradations.

ID: 193
Abstract
Keywords: fMRI, electrocorticography, intracranial EEG, natural sounds, human auditory cortex

Invasive and non-invasive human brain recordings to natural sounds are predictable from a shared subspace common across brain regions and participants

Yongtian Ou[1], Mara Wolter[2], Giorgia Cantisani[3], Sam Norman-Haignere[4,5]
[1]University of Minnesota, Department of Psychology; [2]Centre for Integrative Neuroscience Tübingen; [3]École normale supérieure, PSL; [4]Department of Biostatistics & Computational Biology, Department of Neuroscience, University of Rochester Medical Center, Rochester, USA; [5]Department of Brain & Cognitive Sciences, Department of Biomedical Engineering, University of Rochester, Rochester, USA; giorgia.cantisani@ens.psl.eu

Invasive recordings from the human brain have unparalleled spatiotemporal precision, critical for studying uniquely human functions such as speech and music perception, but have limited spatial coverage and can only be conducted in a small number of specialized patient populations (e.g., epilepsy patients). Integrating invasive recordings with non-invasive neuroimaging methods such as fMRI would thus be potentially transformative. Here, we test the hypothesis that human cortical responses to natural sounds reflect a common, low-dimensional subspace that predicts the vast majority of the response variance across brain regions, participants, and recording modalities.

Using a large dataset of intracranial recordings from 27 patients and fMRI data from 30 participants, we quantified shared variance within and across modalities using cross-participant prediction analyses. Within each modality, neural responses from a single participant could be linearly predicted from those in other participants, explaining 81% of the explainable variance for intracranial recordings and 80% for fMRI. This suggests a substantial fraction of neural responses is common across participants. Cross-modality analyses further revealed that time-averaged intracranial recordings could accurately predict 73% of the explainable variance in fMRI, which is only slightly lower than that of within-modality fMRI prediction (80%). Prediction accuracy was particularly high in the superior temporal gyrus, where intracranial coverage was dense.

Together, these findings demonstrate that most of the human cortical response to natural sounds is predictable by a common subspace that generalizes across participants and recording modalities. This work lays the foundation for integrating invasive and non-invasive human brain responses from non-overlapping participants by inferring a shared subspace common across both modalities.

ID: 194
Abstract
Keywords: speech and music cognition, intracranial EEG, spectrotemporal modulations, auditory representations, superior temporal gyrus

Decoding speech and music from intracranial recordings: evidence for domain-general representations of sound in the human brain

Jérémie Ginzburg[1,2,3,4], Émilie Cloutier Debaque[2,3,4], Arthur Borderie[2,3,4], Laurence Martineau[4], Paule Lessard Bonaventure[4], Robert J Zatorre[1,2], Philippe Albouy[2,3,4]

[1]Montreal Neurological Institute, McGill University, Montreal, Canada; [2]Centre for Research in Brain, Language and Music, Montreal, Canada; [3]CERVO research center, Laval University, Québec, Canada; [4]Centre Hospitalier Universitaire, Laval University, Québec, Canada; jeremie.ginzburg@mcgill.ca

Speech and music represent the most complex sounds that humans use to convey information. Recent studies indicate that the human brain is finely tuned to encode the temporal and spectral features of sounds that are crucial for processing speech and music. However, a key debate persists regarding how the human brain represents these perceptual objects: does it rely on category-based encoding (domain-specific view) or physical sound features (domain-general view)? To investigate this question, we acquired stereo-electroencephalography (sEEG) recordings in eleven patients affected by focal drug-resistant epilepsy with electrodes implanted into auditory areas as they listened to a naturalistic stimulus (the audio track of a famous movie) containing both speech and music. Using logistic regression on behavioral data from 19 healthy participants, we first showed that music categorization relied on high spectral/low temporal modulations, while speech categorization relied on low spectral/high temporal modulations. We then applied ridge regression to sEEG data and found that electrodes in bilateral superior temporal gyri were sensitive to these spectrotemporal modulations (STM). Crucially, using the brain-predicted STM features from these electrodes, we could accurately predict speech/music categorization, demonstrating that STM features encoded in brain activity were sufficient for perceptual classification. Strikingly high correlations ($0.77 < \rho < 0.91$, all corrected $p < .001$) were observed between the STM features reconstructed from brain activity and those of the original audio signal most relevant for behavioral categorization. These correlations also revealed a lateralized dissociation: music categorization was best predicted by activity in the right superior temporal gyrus, while speech categorization was better predicted by activity in the left superior temporal gyrus. These findings suggest that the brain's categorization of speech and music is primarily driven by the spectrotemporal structure of the sound itself, providing robust support for the domain-general view that higher-order auditory representations are shaped by acoustic properties rather than by categorical distinctions. Moreover, the observed lateralization aligns with well-established asymmetries in spectrotemporal processing, with the left hemisphere favoring temporal modulations and the right hemisphere being more sensitive to fine spectral modulations. To further investigate these mechanisms, a complementary study using a similar naturalistic experimental approach with high-field fMRI is currently underway, aiming to explore spectrotemporal processing in the subcortical auditory pathway—an aspect of auditory processing that has never been directly addressed before.

ID: 195

Abstract

Keywords: intracranial, human, single neurons, heschl's gyrus, neurophysiology

Single neuron responses to simple sounds in human auditory cortex and insula during passive listening

Joel I Berger[1], Hiroto Kawasaki[1], Matthew I Banks[2], Sukhbinder Kumar[1], Matthew A Howard[1,3], Kirill V Nourski[1,3]

[1]Department of Neurosurgery, University of Iowa Carver College of Medicine, Iowa City, Iowa 52242, USA; [2]Department of Anesthesiology, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin 53792, USA; [3]Iowa Neuroscience Institute, The University of Iowa, Iowa City 52242, USA; joel-berger@uiowa.edu

There are a limited number of previous studies suggesting that posterior insula is sensitive to sounds. However, these studies almost always involved behaviorally-relevant stimuli. Further, auditory response properties of the insula have not previously been studied on the single-neuron level, while reports of single neurons in human auditory cortex are extremely rare, with only a few published studies to date. Here, we provide the first report of human single neuron data recorded from the insula and provide comparative data from the adjacent primary auditory cortex, recorded intracranially in human participants during passive listening.

Participants were neurosurgical patients undergoing intracranial-EEG monitoring for medically-intractable epilepsy. Stimuli were click trains and pure tones (0.25-8 kHz). Over 330 single neurons were recorded in 11 participants.

In auditory cortex, 84% of all recorded neurons were click-responsive and 95% were responsive to tones. Remarkably, almost a third of neurons in posterior insula and a smaller subset in anterior insula also responded to clicks and pure tones. Responsive neurons were distributed throughout posterior and anterior insula and showed preferred frequency tuning, as did those in auditory cortex. Onset latencies in the insula were similar to those in the primary auditory cortex but response durations were significantly shorter.

The data add to the highly limited literature of human single neuron auditory cortex recordings and show that insula neurons respond to auditory stimuli even in non-behaviorally relevant contexts. The results change our understanding of the insula cortex in the context of audition, suggesting that processing basic auditory stimuli is an important integrative function of insular cortex, which may result from direct connections from auditory thalamus.

ID: 196
Abstract
Keywords: Spatial audio, Change detection, Electroencephalography, Psychophysics, Hearing-in-noise

Assessing Behavioral and Neural Correlates of Change Detection in Spatialized Acoustic Scenes

Katarina C. Poole[1], Drew Cappotto[2], Vincent Martin[1,3], Jakub Sztandera[2], Maria Chait[2], Lorenzo Picinali[1], Martha Shiell[4]
[1]Dyson School of Design Engineering, Imperial College London, United Kingdom; [2]Ear Institute, University College London, United Kingdom; [3]ENTPE, Ecole Centrale de Lyon, CNRS, LTDS, UMR5513, France; [4]Eriksholm Research Centre, Snekkersten, Denmark; katarina.poole@imperial.ac.uk

Detecting changes in complex auditory environments is essential for navigating the world, yet the neural processes supporting this ability are not fully understood. In this study, we explored how the presence and location of sound sources influence both active auditory change detection and the neural signatures of passive change detection. To maintain spatial acoustic cues that were perceptually separable whilst eliminating semantic effects and minimising contextual effects, we used naturalistic amplitude envelopes mapped onto broadband synthetic carriers, presented via a spherical loudspeaker array. We tested participants behaviourally whilst they detected new appearing sound sources in multi-source auditory scenes that were either spatialised or non-spatialised. In the passive listening experiments, participants were given a visual decoy task while neural data were collected via electroencephalography (EEG) during exposure to unattended spatialised scenes and added sources.

Across the two behavioural experiments (N = 21 each), spatialised audio improved change detection accuracy relative to the non-spatialised scenes. However, for the spatialised condition only, performance declined as the number of sources increased and higher high-frequency thresholds also impaired performance. When manipulating the location of the appearing sound source, we identified slower reaction times for changes occurring behind or above the listener, which was exacerbated with a higher number of concurrent sources. EEG experiments using both real and phantom sound source locations (N = 32 and N = 30) revealed reliable neural responses to changes in the auditory scene. Nonetheless, no significant modulation was observed based on the spatial location of the added sound source.

Together, these results highlight the nuanced contributions of spatial cues and scene complexity in auditory change detection, offering new perspectives on how listeners monitor spatial and dynamic acoustic environments. Importantly, these experiments lay the groundwork for future investigations into how individuals perceive and respond to spatialised non-target auditory events and can help evaluate how well current hearing assistive technologies support users' awareness of surrounding soundscapes.

ID: 197

Abstract

Keywords: ASSR, ECoG, gamma oscillation

Gamma-band auditory steady-state responses (ASSR) are altered differently in the frontal and temporal cortices in awake monkey

Mariko Tada[1], Yuki Suda[2], Takanori Uka[2]

[1]Juntendo University Graduate School of Medicine, Japan; [2]University of Yamanashi, Japan; mariko.tada@juntendo.ac.jp

Gamma-band auditory steady-state responses (ASSR) have attracted attention for its relationship with neuropsychiatric disorders such as schizophrenia. In clinical studies, gamma-band ASSR is reduced in patients with schizophrenia as well as other auditory event-related potentials such as MMN. Previously we found that gamma-band ASSRs are observed in wide regions of the cerebral cortex (not only in the auditory cortex but also in the frontal and parietal cortices) using electrocorticography (ECoG) in refractory epilepsy patients. However, the reasons why they are distributed in the brain in such broad areas are not well understood. Using high density ECoG measurements in macaque monkeys (n=3), we employed ASSR, one of the most robust measures of evoked oscillations, to investigate the cortical distribution. The electrode grid was fabricated on a flexible parylene-C film using micro-electro-mechanical systems. ECoG electrodes covered the auditory cortex including the core, belt and parabelt regions, as well as frontal cortices including the frontal pole, dorsal and lateral prefrontal cortex and the orbitofrontal cortex. We found robust gamma-band and beta-band ASSRs in a wide region of the cerebral cortex, which is consistent with our human data. We further tested vulnerability of ASSR to low-dose administration of ketamine. Ketamine is an NMDA receptor antagonist that induces psychotic symptoms in humans, and is thought to be an acute model of schizophrenia. We found that ketamine reduces gamma-band oscillatory responses in the temporal cortex but increases them in the lateral prefrontal cortex. Moreover, beta-band oscillatory responses were enhanced in all cortical area. These results reveal a differential cortical profile of oscillatory responses in the macaque monkey, implying that the local circuits that generate gamma oscillation vary in components according to the cortical area involved.

ID: 198
Abstract
Keywords: auditory cortex, striatum, dopamine, local field potentials, passive listening

Dopamine-Dependent Modulation of Auditory Evoked Potentials in the Striatum During Passive Sound Processing

Andreas L Schulz[1], Michael T Lippert[1], Frank W Ohl[1,2,3]
[1]Leibniz Institute for Neurobiology, Magdeburg, Germany; [2]Otto-von-Guericke University,Magdeburg, Germany; [3]Center for Behavioral Brain Sciences CBBS, Magdeburg, Germany; schulza@lin-magdeburg.de

The cortico-striatal pathway is involved in the transformation of auditory representations into decision making and consequently into motor commands. Learning of an auditory discrimination task potentiates cortico-striatal synapses. It remains unknown so far, whether dopamine is necessary for plastic changes of these synapses. To assess the role of dopamine on these synapses we employed Mongolian gerbils in a passive listening task in which presentation of frequency modulated tones was paired with optogenetic stimulation of the ventral tegmental area (VTA). We recorded local field potentials in two different sites in the striatum, in the posterior and the dorsal lateral striatum. The amplitude of the auditory evoked potentials (AEP) in the striatum was increased after VTA stimulation in the dorsal striatum but not in the posterior striatum. Systemic application of a dopamine antagonist (SCH-23390) prevented the increase of the AEP amplitude and sometimes caused it to be even smaller after VTA stimulation. AEP amplitudes were not affected without VTA stimulation. In summary, dopamine seems to be required for the increase of striatal AEP amplitude and consequently for potentiation of auditory striatal synapses, however in the current study we did not distinguish between alternative possible pathways, like cortico-striatal or thalamo-striatal connections.

ID: 199
Abstract
Keywords: auditory adaptation, interval, dynamic range, MEG, efficient coding

Adaptive Encoding of Temporal Intervals in Human Auditory Cortex

Honghua Chen[1], Jing Wang[1], Jiaxin Gao[1], Nai Ding[1,2]
[1]Key Laboratory for Biomedical Engineering of Ministry of Education, College of Biomedical Engineering and Instrument Sciences, Zhejiang University, Hangzhou, China; [2]The State key Lab of Brain-Machine Intelligence; The MOE Frontier Science Center for Brain Science & Brain-machine Integration, Zhejiang University, Hangzhou, China; honghuachen@zju.edu.cn

Encoding time intervals is critical for auditory perception. Here, we investigated how the human auditory cortex encode the time intervals between events in different contexts. Using magnetoencephalography (MEG), we recorded neural responses from participants exposed to syllable sequences of different rhythmic properties. The mean inter-syllable-onset interval was either 333 ms or 200 ms and the coefficient of variance (CV) of the inter-syllable-onset interval ranged from 0.25 to 1.0. Analysis revealed that the M100 component reliably encoded interval information across all conditions, which was characterized by a interval tuning curve. Crucially, the interval tuning curve depended on both the mean and CV of the inter-syllable-onset intervals. Source localization traced the rhythm-dependent interval encoding patterns to mainly the right auditory cortex. The rhythm-dependent interval coding mechanism may enhance interval encoding precision in context.

ID: 200
Abstract
Keywords: speech production, synchronization

Individual rhythms in solo and joint speech

Max Austin Wolpert, Nai Ding
Key Laboratory for Biomedical Engineering of Ministry of Education, College of Biomedical Engineering and Instrument Sciences, Zhejiang University, Hangzhou, China; max.wolpert@zju.edu.cn

INTRODUCTION

Individuals exhibit stable rate preferences for motor activity, like finger tapping or music performance, and these preferences also impact auditory-motor synchronization (Scheurich, Zamm, & Palmer, 2018). We extend consideration to speech production, a highly adaptable task with layered rhythmic structure: pulse (syllabic timing) and meter (prosodic phrasing) (Fitch, 2013). To balance between these two rhythms, speakers may modulate syllabic rate to preserve prosodic timing across longer phrases (Jun 2003).

METHODS

We use a self-paced and joint-reading paradigm to examine:

1. Whether individuals show stable self-paced speech rate preferences and whether prosodic structure affects these rates;
2. Whether individual rate preferences influence synchronization in joint speech.

Ninety participants first completed a self-paced reading task. Based on their average syllable inter-onset interval (IOI), we invited a subset of 32 to perform synchronized joint reading. Each participant read with one rate-matched and one rate-mismatched partner. Stimuli were short sentences of varying lengths (4–6 syllables), and were read in blocks of 10 sentences to promote rhythmic entrainment. We measured speech rhythm via syllable IOIs (including lengthening and pauses) and synchronization via onset asynchrony. The difference in solo IOI rate preferences within each pair was used to predict joint asynchrony.

RESULTS

Self-paced IOIs ranged from 188 to 388 ms (M = 288, SD = 50), showing substantial inter-individual variation. IOIs decreased as sequence length increased (~30 ms from 4 to 6 syllables), consistent with metrical pacing. IOIs at sentence-final positions were about twice as long as in other positions, indicating the prominence of phrase-final lengthening and pause.

In joint reading, rate differences between partners did not significantly predict onset asynchrony, though pairs with closer solo rates showed a small numerical trend toward better synchronization. Regardless of initial preference, pairs tended to converge toward an intermediate rate, with one speaker leading by ~30–40 ms. Asynchrony peaked at sentence onsets and decreased over the course of the sentence, suggesting participants used both sentence-initial and sentence-final boundaries as synchronization cues.

CONCLUSION

These findings suggest that while individuals have preferred speech rates, they flexibly adapt in social contexts. The weak predictive power of solo speech rate for joint synchronization highlights the flexibility of entrainment mechanisms. Compared to music, speech's faster tempos and greater variability may make temporal alignment

less reliant on individual rhythmic defaults and more dependent on adaptive cueing at prosodic boundaries.

ID: 201

Abstract

Keywords: speech;information theory;predictive coding;entropy;intracortical eeg

Differential processing of predictive strength and dispersion in naturalistic speech

Pierre Hieu Guilleminot, Benjamin Morillon

Institut de Neuroscience des Systèmes, France; pierre.guilleminot@univ-amu.fr

Speech comprehension relies on predictive mechanisms that minimize error between expected and incoming auditory input. The neural response to a word is therefore modulated by its surprisal (its likelihood) and by the uncertainty of the context (its overall predictability). This contextual uncertainty is typically quantified using Shannon entropy, which captures the overall unpredictability of a given word based on prior context. However, Shannon entropy is only one member of the broader Rényi entropy family, which enables a distinction between uncertainty due to the strength of dominant predictions and that due to the dispersion across alternatives. Here, we investigated how different measures of uncertainty modulate the neural response to words during naturalistic speech perception. Intracortical recordings from 33 epilepsy patients revealed that the brain processes prediction strength and dispersion in distinct neural populations, with dispersion linked to lower-level auditory areas and strength to more associative regions. Moreover, brain responses in these areas were differentially modulated by their respective uncertainty measures. These findings suggest that the auditory system encodes a richer, more differentiated representation of contextual uncertainty than previously assumed, pointing to a more complex role for predictive processing in speech comprehension.

ID: 202

Abstract

Keywords: neural oscillations, temporal prediction, phase, EEG, attention

A phase-based mechanism of attentional selection in the auditory system

Troby Ka-Yan Lui[1,2], Victorina De Combles De Nayves[3], Céline Cappe[1,2], Benedikt Zoefel[1,2]

[1]Centre de Recherche Cerveau et Cognition (CerCo), CNRS, Toulouse, France; [2]Université de Toulouse, France; [3]Université de Strasbourg, France; benedikt.zoefel@cnrs.fr

Perceptual targets are often easier to detect and process if they occur at a predictable time. This per-ceptual benefit of temporal predictability suggests that neural resources are allocated to specific mo-ments predicted to be most informative, or away from those predicted to be distracting. The sensitivity to sensory information changes with the phase of neural oscillations. An alignment of oscillatory phase according to the predicted time and relevance of expected events has therefore been proposed as a mechanism underlying attentional selection. This hypothesis predicts that opposite phases are aligned to relevant and distracting events so that the former are amplified and the latter suppressed, but such an effect remained to be demonstrated.

We presented 25 human participants with acoustic stimuli that occurred at a predictable or unpredicta-ble time during a pitch-discrimination task. Temporal predictability did not only lead to faster reaction times, it also aligned the phase of beta oscillations in the electroencephalogram (EEG) in anticipation of the stimulus. Crucially, the aligned phase was opposite, depending on whether the stimulus was relevant or a distractor for the task. All effects occurred in the absence of rhythmic stimulation, an important but rarely tested criterion to rule out spurious phase alignment from preceding events in a stimulus sequence.

We therefore conclude that the alignment of oscillatory phase underlies the selection of sensory in-formation in time. We speculate that beta oscillations allocate neural resources in auditory-motor net-works to the expected time of auditory information, facilitating behavioural responses to upcoming events.

ID: 203
Abstract
Keywords: acoustic texture, perception, continuity illusion, intracranial EEG

Neural bases of illusory acoustic texture perception

Alexander J Billig[1], Joel I Berger[2], Richard McWalter[3], Tobias Teichert[4], Christopher M Garcia[2], Christopher K Kovach[2,5], Christopher I Petkov[2], Hiroto Kawasaki[2], Matthew A Howard[2], Josh H McDermott[3], Timothy D Griffiths[2,6,7]
[1]UCL Ear Institute, University College London; [2]Department of Neurosurgery, University of Iowa; [3]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology; [4]Department of Psychiatry and Bioengineering, University of Pittsburgh; [5]Department of Neurosurgery, University of Nebraska Medical Center; [6]Biosciences Institute, Newcastle University; [7]Department of Imaging Neuroscience, University College London; a.billig@ucl.ac.uk

Many of the sound scenes we encounter consist of a large number of stochastically timed similar events that together form acoustic textures, such as the sound of a rain shower. These can typically be characterised by a relatively sparse set of summary statistics, such as correlations between modulation envelopes across frequency bands (McDermott & Simoncelli, Neuron 2011). Listeners are thought to automatically extract such statistics when perceiving and remembering textures (McDermott et al., Nat Neurosci 2013). Like other sounds, textures can be perceived to continue through sufficiently loud interrupting white noise, even when the texture is physically absent (McWalter & McDermott, Nat Comms 2019). In the case of textures with stable summary statistics, this illusion can last for several seconds. We tested the hypothesis that such persistence draws on neural circuits beyond auditory cortex, including hippocampus. We presented many exemplars of two different textures to ten neurosurgical participants. In each trial, two seconds of texture were followed by two seconds of white noise and then a final second of texture. In a control condition, 200-ms silent gaps were inserted either side of the noise. All participants reported perceiving the illusion in the continuous case, with a significant drop in such reports when gaps were present. Only when the texture was physically present could its identity be decoded from single-neuron firing and high-gamma power in auditory cortex. In contrast, decoders using the raw local field potential or theta power across multiple electrodes were often able to identify not only the physically presented texture, but also the texture perceived (but physically absent) during the noise. Sites beyond primary auditory cortex, including in superior temporal sulcus, temporal pole, planum temporale and the medial temporal lobe, carried the greatest weight in these decoders. In the majority of participants, this decodability dropped to chance level when silent gaps buttressed the noise to suppress the illusion. We continue to study the extent to which the extraction and persistence of summary statistics through several seconds of interruption draws on brain areas with longer intrinsic timescales than primary auditory cortex, and interactions between these regions.

ID: 205
Abstract
Keywords: cortical stimulation, ECoG, AEP, minipig

Cortical Stimulation of the Auditory Cortex to Restore Hearing: Validation In-Vivo in Minipigs

Margaux Roulet[1,2], Florian Fallegger[3], Lukas Anschuetz[1,2], Stéphanie Lacour[3], Alix Trouillet[1,2,3]
[1]Translational Auditory Research Laboratory, Department of Surgery, Lausanne University Hospital (CHUV) and University of Lausanne (UNIL), Lausanne, Switzerland; [2]The Sense Innovation and Research Center, Lausanne, Switzerland; [3]Laboratory for Soft Bioelectronic Interfaces, Neuro-X Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Geneva, Switzerland; margaux.roulet@chuv.ch

The auditory cortex (AC) is a promising target for novel auditory prosthesis. Inspired by cortical stimulation successes in motor and other sensory perception restoration, an auditory cortex prosthesis (ACP) could offer a personalized solution for patients with hearing loss or tinnitus. Targeting the AC has the advantage of being applicable to patients with auditory disorders affected in their peripheral and central pathways. In addition, the large surface area of the AC opens the way for more advanced devices with enhanced spatiotemporal resolution and stimulation generators that dynamically adapt to neural responses.

We aim to develop a cutting-edge therapy for patients with hearing loss who cannot benefit from available technology. Towards this path, we built a minimally invasive surface ACP using thin-film technology and assessed its functionnality in-vivo in minipigs. We explored cortical stimulation as a mean to activate the auditory pathway and elicit auditory percepts: we applied electrical cortical stimulation on one hemisphere and recorded the contralateral activation patterns. In parallel, we recorded brain activity upon acoustic stimulation to serve as functional reference.

Minipigs were implanted at the surface of the auditory cortices with ECoG devices. First, we recorded brain activity upon acoustic stimulation using tone bursts at frequencies ranging from 200 Hz to 20 kHz at 70 dB. We then substituted auditory stimulation by electrical stimulation which was applied through the ACP using an amplitude ramping protocol and varying monopolar and bipolar stimulation paradigms. Event-related potentials were extracted and averaged over 100 repetitions. We analyzed the auditory evoked potential induced by acoustic stimulation (AEP) and electrical stimulation (eAEP) and performed a correlation analysis to assess the relationship in brain activity elicited by a specific stimulation paradigm and acoustic frequencies. Ramping curves across stimulation amplitudes, and in the presence of concurrent white noise was also evaluated.

We recorded AEP maps consistent with the tonotopic organization of the AC in both hemispheres under acoustic stimulation. Electrical stimulation of the AC elicited brain responses that matched these maps. Concurrent white noise significantly attenuated eAEPs, mirroring its effect on AEPs and therefore validating the auditory percept of the electrical stimulus. More importantly, we demonstrated that monopolar and bipolar stimulation paradigms are spatially selective enough to elicit distinct brain responses. We correlated eAEPs and AEPs maps and discover frequency selectivity of cortical stimulation.

Results suggest that cortical stimulation can reproduce natural auditory percept. Next addressed challenge will include optimizing the precision and targeting of electrodes, in-depth understanding of long-term effects of cortical stimulation on brain plasticity and finally, encoding of complex sound stimuli such as music.

ID: 206
Abstract
Keywords: speech, music, rhythm, modulation spectrum, vocalization

Rhythm of speech, music, and other vocalizations

Jing Wang, Wei Liu, Nai Ding
College of Biomedical Engineering and Instrument Sciences, Zhejiang Univeristy, China; ding_nai@zju.edu.cn

Speech and music both have their own rhythmic characteristics. Recent studies consistently demonstrate that the dominant acoustic rhythm of speech is centered around 4-5 Hz while the dominant acoustic rhythm of music, including songs, is centered below 2 Hz. Here, we investigate whether the rhythm of human vocalization consistently falls into a dichotomy, i.e., either speech- or music-like, or whether the rhythm of vocalization can form a continuum. By analyzing of a wide range of human vocalizations (including crying, laughing, commands, literature chanting, city cries, and various speech-music art forms) across ages (from infancy to adult) and across cultures, we observe that speech- and music-like rhythms are reliable categories but diverse rhythmic conditions exist between them. Together, these findings challenge the strict dichotomy between speech and music and highlight a shared temporal architecture that is graded across contexts and emerges with development.

ID: 207
Abstract
Keywords: binaural hearing, neural tracking, Cochlear implant

Neural responses of binaural hearing in simulated Cochlear Implant users

Juan-Daniel Galeano-Otálvaro[1], Tom Francart[1], Benjamin Dieudonné[1,2], Jan Wouters[1]
[1]ExpORL, KU Leuven, Belgium; [2]BrainPhonics BV, Belgium;
juandaniel.galeanootalvaro@kuleuven.be

Cochlear implants (CIs) restore access to sound and enable speech comprehension in individuals with severe-to-profound hearing loss. However, CI users still experience considerable difficulty understanding speech in noisy environments, especially compared to individuals with normal hearing (NH). A major contributing factor is the inadequate representation of binaural cues—specifically interaural time differences (ITDs) and interaural level differences (ILDs)—which are vital for spatial hearing and separating speech from background noise. Users of bilateral cochlear implants (BiCIs) often struggle because each implant processes sounds independently, which disrupts binaural cue transmission, impairs sound localization, and reduces speech-in-noise performance.

Traditionally, binaural hearing is evaluated through subjective behavioural measures such as binaural masking level differences (BMLDs). However, CI users tend to show high variability in these outcomes and often only derive limited benefit from binaural input. This highlights the need to better understand the neural basis of binaural processing and how it could be harnessed to enhance CI outcomes. Recent research has started addressing this by using electroencephalography (EEG) to objectively assess binaural benefits in speech understanding. EEG provides a way to link neural responses to binaural cues with perceptual performance, offering novel insights into the mechanisms of binaural hearing and individual differences.

In this study, we explore masking release and neural markers of binaural processing in NH listeners using simulated BiCI hearing. Binaural hearing performance is evaluated with speech and speech-like signals. The stimuli are pre-processed to mimic the reduced temporal fine structure typical of CI sound processing. Behavioral thresholds for masking release are assessed by comparing responses in a diotic condition to dichotic conditions containing interaural cues. Neural responses will be recorded using EEG.

Neural tracking differences between masked and unmasked conditions will be assessed for both types of stimuli, and evoked auditory response amplitudes elicited by the envelope stimuli will be analyzed. We hypothesize that behavioral masking release will correspond to enhanced neural tracking and stronger neural response amplitudes under conditions that exploit binaural cues. Finally, as different modulation rates will be used, we expect them to elicit activity in different regions along the auditory pathway. Our findings aim to bridge the gap between perceptual and electrophysiological assessments of binaural hearing. The results will provide a foundation for understanding how neural processing of binaural cues relates to auditory performance, allow to measure binaural processing objectively in CI users, and develop stimulation strategies.

This work was supported by the Marie Skłodowska-Curie Actions (MSCA) of the European Union, grant No. 101120054.

ID: 208

Abstract

Keywords: Auditory processing; Audio-tactile processing; Phase effect; Top-down modulation; Electroencephalography;

Top-down attention modulates phase effects on audio-tactile processing

Xueying Fu[1,2], Lars Riecke[1]

[1]Maastricht University, the Netherlands; [2]China; xueyingfu.psy@gmail.com

Multisensory processing is necessary in our daily life, which is modulated by the temporal alignment of multiple sensory inputs. However, it remains poorly understood if and how selective attention interacts with this temporal alignment to further modulate audio-tactile processing. This study aimed to explore how selective attention and temporal alignment influence cortical responses to rhythmic audio-tactile streams. Participants were exposed to periodic auditory tones embedded in continuous background noise, either presented alone or paired with fluctuating tactile stimulation that was either in-phase or anti-phase with the tones. Selective attention was manipulated by instructing participants to perform an auditory detection task focusing on either the tones or the background noise. Electroencephalography recordings revealed that anti-phase audio-tactile inputs, compared to auditory-only inputs, elicited enhanced cortical steady-state responses and phase-locking to the tones. Importantly, this effect was further enhanced when participants paid attention to the tone, potentially reflecting the resolution of sensory competition among the simultaneous auditory and tactile inputs. In-phase tactile inputs elicited non-significant increases in tone processing regardless of attention. In sum, these findings underscore the dynamic interplay between bottom-up temporal alignment and top-down attentional modulations in audio-tactile processing, providing insights into how the brain integrates audio-tactile information.

ID: 209

Abstract

Keywords: Spiking Neural Networks, Auditory Processing, Rhythmic Perception, Spiking Encoding

A spiking neural network for investigating auditory rhythm processing

Rodrigo Manríquez[1,2], Sonja A. Kotz[2,3], Andrea Ravignani[4,5], Bart de Boer[1]

[1]Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussel, Belgium; [2]Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands; [3]Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; [4]Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy; [5]Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music Aarhus/Aalborg, Denmark; rmanriquez@ai.vub.ac.be

While artificial intelligence and deep learning models have gained significant attention in neuroscience, there is still a need for more biologically realistic models to capture the dynamics of specific brain functions. Spiking Neural Networks (SNNs) may fill this gap by using the precise timing of spikes as the main mechanism for information processing, offering greater biological plausibility. In this study, we put forward a spiking neural network framework for auditory rhythm processing, focusing on its detection and representation.

Our approach involves encoding auditory signals into spike trains, using a biologically inspired subcortical model of sound processing. This model simulates peripheral auditory functions, particularly the auditory transduction that occurs at the cochlear level, by reproducing auditory nerve responses tuned to specific characteristic frequencies. For each characteristic frequency, a spike train is generated. In this way, raw acoustic waveforms are converted into a temporally precise spiking representation that preserves key temporal features of the input. The encoded spiking data is then processed using a spiking autoencoder, a neural architecture designed to learn efficient representations of the input. The autoencoder is trained to reconstruct the amplitude envelope of the acoustic signal at its output layer, effectively capturing rhythmic and amplitude modulations present in the original sound.

Through our simulations, we can demonstrate that when the network is trained on isochronous rhythmic sequences, i.e. acoustic sequences where time intervals between successive events have the same duration, emergent rhythmic patterns materialize in the latent representations learned by the spiking autoencoder. The network learns the timing of onsets and develops predictive capabilities, allowing for the anticipation of subsequent rhythmic events. This sensitivity reflects a form of temporal expectation encoded within the spike-based architecture.

To further investigate the network's ability to encode rhythmic structures, we evaluate its performance with alternative rhythmic paradigms, such as missing beats and alternating high and low-amplitude pulses. These simulations can reveal the nature of timing representations learned by the network, emphasising how it relies on temporal features of the input to anticipate subsequent pulses. Moreover, our findings support the hypothesis that rhythm encoding can arise from purely spike-based processing, also reinforcing the biological plausibility of SNNs in auditory neuroscience research.

ID: 210
Abstract
Keywords: Puberty, Adolescence, Speech perception in noise, EEG, functional connectivity, cortical tracking of speech in noise, cognitive development, neurodevelopment

Puberty-related electrophysiological changes underlying speech perception in noise

Francesca Cavicchiolo[1,2,3], Marlies Gillis[1,2], Joseph Francois Johnson[1,2,3], Marta Puertollano[1,2,3], Luna Prud'homme[1,2,3], Vincent Wens[1,2,4,5], Xavier De Tiège[1,2,3,4,5], Axelle Calcus[1,2,3]

[1]Université libre de Bruxelles (ULB); [2]UNI – ULB Neuroscience Institute; [3]CRCN – Center for Research in Cognition and Neuroscience - ULB; [4]H.U.B. – ULB University Hospital; [5]Laboratoire de Neuroanatomie et Neuroimagerie translationnelles (LN2T) - ULB; francesca.cavicchiolo@ulb.be

The maturational process of the human brain and behavior towards adulthood is associated with the improvement of cognitive and perceptual skills required to succeed in complex social environments. Among these, the ability to perceive speech signals in the presence of noise or competing speakers is crucial in teenagers' daily lives. The development of speech processing in noise (SIN) extends into adolescence. Yet the neurobiological mechanisms underlying SIN remain poorly understood, but likely stem from changes at the level of the endocrine and nervous systems.

As part of the SensationaHL pubertal development cohort collection, this study investigates the effects of puberty onset on the electrophysiological changes underlying SIN development. By means of a combination of behavioral tasks and high-density electroencephalography, we are testing the hypothesis that puberty onset triggers a period of heightened neuroplasticity for SIN improvement. To this aim, we are comparing age- and sex-matched pre- and early-pubertal adolescents on functional connectivity both at rest and during a SIN task, cortical tracking of speech in noise and natural speech oddball evoked responses. Furthermore, we are investigating the relationship between electrophysiological and behavioral (SIN and cognition) development.

We will present preliminary results from a subset of participants. We expect to observe region-specific between-groups differences in functional connectivity, as well as differences in cortical tracking and in the amplitude and latency of oddball evoked-responses. Finally, we expect to observe a relationship between neural and behavioral measures.

This study will further our understanding of the neurobiology of adolescent development. Furthermore, it holds practical implications: the emphasis on speech-in-noise perception addresses a real-world challenge with potential effects on academic performance and social interactions during this critical developmental phase.

ID: 211
Abstract
Keywords: corticocollicular, learning, two-photon

Auditory learning modulates sound and choice selectivity of layer 5 corticocollicular neurons

Kira M.A. Andrea, Jan J. Hirtz
Animal Physiology, Department of Biology, RPTU University of Kaiserslautern, Germany; hirtz@bio.uni-kl.de

Corticocollicular (CC) projections of the auditory cortex (AC) are crucially involved in auditory-related learning and predictive coding. We here investigated activity of CC neurons in mouse AC layer 5 using two-photon imaging during a Go/No-go learning paradigm spanning two weeks. Animals were trained to lick a waterspout in response to a target tone, while withholding licking when a foil tone was presented. During the training process, the selectivity of CC neurons shifted towards the foil tone. However, when taking decision making of the animal into account, failing to lick in response to the target sound (miss response) elicited strongest activity compared to other behavioral situations. In contrast, preliminary data obtained from intratelencephalic L5 AC neurons demonstrate a shift of activity towards the target tone and behavioral decisions other than the miss response. We prior showed that layer 5 CC neurons display only a subtle topographic organization when compared to other layer 5 neuron types (Schmitt et al., 2023, Front Neural Circuit). While these data were acquired under passive listening conditions, we also could not detect any sharpening of topographical organization of sound-evoked activity during the learning process, suggesting that weak tonotopy of CC neurons is not related to the behavioral relevance of sounds. Our data further highlight the importance of layer 5 CC neurons in auditory learning, expressed in both sound- and behavior-related activity shifts which are distinct from other layer 5 AC neuron types.

ID: 212
Abstract
Keywords: Sound localisation, auditory cortex, ferret

Neural Representations of Auditory Space During Sound Localisation

Stefan Zigic, Amit Khandhadia, Nadia Aghili Kordmahale, Jennifer Bizley
University College of London, United Kingdom; s.zigic@ucl.ac.uk

Understanding how the brain constructs representations of auditory space is crucial for deciphering how animals navigate complex acoustic environments. Traditional studies have primarily investigated spatial hearing using head-fixed subjects with sounds presented from limited, peripheral locations. In contrast, research involving freely moving animals offers a more comprehensive perspective on auditory spatial representation. Notably, a study by Town et al. (2017) recorded neural activity in the auditory cortex of freely moving ferrets and found that the majority of spatially tuned neurons encoded sound source locations relative to the head, indicating an egocentric reference frame. However, a subset of neurons exhibited tuning to sound locations independent of the animal's position, suggesting the presence of allocentric, or world-centered, representations. To further investigate the neural representation of space, we have trained ferrets to navigate a 2x4m arena with an acoustically transparent mesh floor, under which are 40 individually controlled speakers forming a speaker grid. Ferrets were trained to begin trials at a start location and navigate towards sound sources. On each trial, two speakers were randomly selected and once a trial was initiated a speaker produced an irregular sequence of 100ms noise bursts at 150–500ms intervals. Animals had to navigate through a 22cm diameter zone based on the centre of the speaker, and following the detection of a traversal, a second speaker would activate. The subject would then have to navigate to the location of the second sound, pausing within the detection zone for 1800ms, which was followed by the delivery of a water reward at reward zone. This task was controlled by DeepLabCut Live, which allowed the animal's location to be tracked in real time. The ferrets rapidly learned to perform the task, and current experiments are limiting the availability of localization cues by utilizing band-limited stimuli. One animal has been recently implanted bilaterally with Neuropixel 2.0s in the auditory cortex, with a trajectory that includes hippocampus. We aim to determine the auditory spatial representation by integrating neural data with positional and angular data of the animal, within the arena and relative to active speakers. This approach will allow us to probe how auditory space is encoded during active listening and real-world navigation.

ID: 213

Abstract

Keywords: TRN, ultra-high field MRI, quantitative T2*, human brain

Mapping the human thalamic reticular nucleus (TRN) by means of an ultra-high field 7T MRI: Toward a probabilistic atlas

Zuzanna Alicja Kotwicka[1,2], Faruk Gulban[3,4], Logan Dowdle[3], Ryszard Auksztulewicz[2], Michelle Moerel[1,3]

[1]Maastricht Centre for Systems Biology and Bioinformatics (MaCSBio), Faculty of Science and Engineering, Maastricht University, Maastricht, the Netherlands; [2]Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands; [3]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands; [4]Brain Innovation B.V., Maastricht, The Netherlands; z.kotwicka@student.maastrichtuniversity.nl

Auditory attention, the ability to selectively process relevant sounds while filtering out distractions, is a critical cognitive function that relies on the regulation of sensory information flow. The thalamic reticular nucleus (TRN), a thin layer of inhibitory GABAergic neurons surrounding the thalamus (1), has emerged as a key structure in attentional gating. While animal studies have implicated the TRN in modulating ascending auditory signals (2), its role in the human brain remains poorly understood due to its small size and deep anatomical location, which present challenges for studying the TRN with conventional non-invasive imaging techniques. In the present study, we leverage the high spatial resolution provided by ultra-high field (UHF) magnetic resonance imaging (MRI) at 7 Tesla to assess the feasibility of reliably visualizing the TRN in vivo.

We acquired quantitative T1 and T2* partial brain datasets targeting subcortical regions, at 0.35 × 0.35 × 0.35 mm3 voxel resolution (3) from 11 healthy young subjects. The data underwent thermal noise reduction using tNORDIC (4), followed by preprocessing (3) and smoothing with the Segmentator Python package (5,6). Our preliminary findings show that the TRN is visible in 7 out of 11 subjects on the quantitative T2* images, demonstrating the feasibility of in vivo visualization at 7 Tesla. Datasets of the remaining 4 subjects were of lower quality, likely due to the incomplete nature of these datasets or motion artifacts.

As next steps, we plan to manually segment the TRN, followed by quantitative assessment of inter- and intrasubject variability in its volume, surface area, and thickness. Our ultimate goal is to provide a probabilistic atlas of the human TRN, providing a foundation for future studies of auditory attention.

1 Li, Y. et al. Distinct subnetworks of the thalamic reticular nucleus. Nature 583, 819-824 (2020). https://doi.org/10.1038/s41586-020-2504-5

2 Homma, N. Y. et al. A Role for Auditory Corticothalamic Feedback in the Perception of Complex Sounds. J Neurosci 37, 6149–6161 (2017). https://doi.org/10.1523/JNEUROSCI.0397-17.2017

3 Gulban, O. F. et al. Mesoscopic in vivo human T2* dataset acquired using quantitative MRI at 7 Tesla. Neuroimage 264, 119733 (2022). https://doi.org/10.1016/j.neuroimage.2022.119733

4 Ponticorvo, S. et al. Reducing thermal noise in high-resolution quantitative magnetic resonance imaging rotating frame relaxation mapping of the human brain at 3 T. NMR in Biomedicine 37 (2024). https://doi.org/10.1002/nbm.5228

5 Gulban, O. F., Schneider, M., Marquardt, I., Haast, R. A. M. & De Martino, F. A scalable method to improve gray matter segmentation at ultra high field MRI. PLOS ONE 13, e0198335 (2018). https://doi.org/10.1371/journal.pone.0198335

6 Segmentator v. 1.5.3 (Zenodo, 2019). https://doi.org/10.5281/zenodo.2601899

ID: 214
Abstract
Keywords: musical behavior, electroencephalography, rhythm perception and production, perceptual categorization, cross-cultural comparison

Investigating culture-driven neural plasticity in rhythm categorization

Diana-Ioana Mares[1], Tomas Lenc[1,2], Francesca M. Barbero[1], Rainer Polak[3,4], Nori Jacoby[5,6], Sylvie Nozaradan[1,7]

[1]Institute of Neuroscience (IoNS), University of Louvain (UCLouvain), 1348 Louvain-la-Neuve, Belgium; [2]Basque Center on Cognition, Brain and Language (BCBL), Donostia-San Sebastian, Spain; [3]RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Norway; [4]Department of Musicology, University of Oslo, Norway; [5]Department of Psychology, Cornell University, Ithaca, NY 14853, United States; [6]Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt am Main, Germany; [7]International Laboratory for Brain, Music and Sound Research (BRAMS), Montreal, Canada; diana-ioana.mares@uclouvain.be

Humans effortlessly learn, recognize, and move to musical rhythms. This ability builds on a categorization process that maps the vast diversity of external rhythms onto a limited repertoire of internal categories. Behavioral studies have shown that while certain rhythm categories are universal, others differ across cultures, influenced by the prevalent rhythmic patterns in local musical traditions. However, the neural mechanisms behind the cross-cultural diversity of rhythm categorization remain still largely unknown.

Recent research using a combination of electroencephalography (EEG), frequency-tagging, and representational similarity analysis (fRSA) provided evidence for rhythm categories spontaneously emerging in human brain activity during listening to rhythmic sequences in task-free settings. Here, we aim to move beyond these initial findings by exploring how participants with different levels of enculturation respond to both universal and culturally variant rhythms. To this aim, brain activity will be recorded using EEG as participants listen to repeated rhythmic patterns made of three inter-onset intervals. The interval durations will be manipulated over 12 conditions to yield a continuum of stimuli covering the space between two universal rhythms with inter-onset interval ratios of 1:1:1 (isochrony) and 1:1:2 (short-short-long), respectively. Notably, the continuum will also include one culturally specific rhythm, with a 2:2:3 ratio, characteristic of South-Eastern European musical traditions. Familiarity with the rhythms is manipulated by comparing the EEG of participants with vs. without substantial familiarity with South-Eastern European music styles (i.e., South-Eastern European vs. Western European participants). Complementing the EEG data, participants' sensorimotor synchronization (finger tapping) to the same stimuli will be measured to obtain a behavioral index of rhythm categorization.

Data acquisition is currently underway. We expect to find significant categorization in both neural and behavioral responses, irrespective of participants' group. However, we hypothesize that the nature of categorization will differ between the groups, with South-Eastern European participants showing the presence of an additional category corresponding to the culture-specific rhythmic ratio, as compared to Western participants. By highlighting the differences in neural representations across participants differing in long-term experience with the rhythms, these results are expected to provide fundamental insight into the plasticity and automaticity of universal vs. non-universal rhythm categories. More generally, this study aims to pave

the way for future research clarifying the nature of the social and neurobiological processes that support culture-driven plasticity in rhythm perception from birth.

ID: 215
Abstract
Keywords: Sound Textures, Summary Statisitcs, Auditory Cortex, Hippocampus, fMRI

Auditory Cortex Representation of Naturalness in Sound Textures

Yousef Mohammadi[1,2], Alexander Billig[3], Richard McWalter[4], Martina Callaghan[2], Timothy Griffiths[1,2]
[1]Biosciences Institute, Newcastle University, UK; [2]Department of Imaging Neuroscience, Univesity College London, UK; [3]UCL Ear Institute, Univesity College London, UK; [4]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology; yousef.mohammadi@newcastle.ac.uk

Natural sound textures, such as rain or crackling fire, are fundamental to our auditory environment. These sounds are defined by time-averaged summary statistics that capture their perceptual identity (McDermott & Simoncelli, 2011). The naturalness of synthetic textures depends on how closely these statistics match real-world textures (Freeman et al., 2013). In this study, using fMRI we examined brain responses to sound textures with parametrically varied summary statistics. This allowed us to probe neural sensitivity to changes in naturalness, revealing graded responses in auditory and non-auditory regions.

Thirty volunteers underwent 3T fMRI while performing an active listening task. MRI images were continuously acquired (TR = 1.1 s, TE = 15.85 ms, 3D sequence, whole-brain acquisition, isotropic voxel size = 3 mm). Sounds were delivered using a bespoke high-fidelity delivery system based on parallel piezoelectric transducers. We measured texture statistics from real-world recordings processed through a biologically inspired auditory model. Sets of synthesized textures that varied in naturalness along a continuum between noise and texture while preserving its general acoustic category were generated. The level of naturalness was controlled by adjusting the proportion of texture to spectrally matched noise on a logarithmic scale, with texture proportions of 0.25, 0.4, 0.63, and 1.0. In each trial, participants listened to two sounds ("reference" then "test") based on different exemplars of the same texture and were required to indicate whether the level of naturalness was the same in each, which occurred in 25% of trials. Four texture categories were used, with stimuli presented in a pseudo-randomized order over five runs (52 trials per run). Univariate analysis assessed voxel-based neural responses to naturalness as a modulatory effect. Additionally, a psychophysiological interaction (PPI) analysis examined whether connectivity between auditory cortex subregions and medial temporal lobe (MTL) regions was modulated by naturalness.

Our results revealed significant ($p < 0.05$, FWE-corrected) bilateral activation in Te1.0, Te1.1, Te1.2, Te2.1, Te2.2, and Te3 in response to increasing naturalness. At an uncorrected threshold ($p < 0.001$), we also observed bilateral entorhinal cortex activation in response to increasing naturalness, but only during test sound presentation. PPI analysis showed a significant ($p < 0.05$, FDR-corrected) negative modulatory effect of naturalness between the right hippocampus (HC) and right higher-order auditory areas (Te1.2, Te2.2, Te3), but only during the perception of the reference sound. In other words, as the sound texture became more noise-like, connectivity between the HC and AC strengthened. These findings suggest that the responses in the auditory cortex may be driven by time-averaged summary statistics that define texture naturalness, while the MTL, particularly the hippocampus, may contribute to top-down mechanisms involved in texture recognition.

ID: 216

Abstract

Keywords: Excitation-inhibition ratio, cortical gradients, auditory cortex, connectivity, large-scale cortical models

Gradients of neural excitation-inhibition ratio across areas of human auditory cortex

Mohsen Alavash[1], Riccardo Catena[1], Björn Herrmann[2], Jonas Obleser[1]

[1]*University of Lübeck, Germany;* [2]*University of Toronto, Canada; m.alavash@uni-luebeck.de*

The excitation-inhibition (E/I) ratio is a fundamental parameter in auditory neuroscience because it shapes the temporal and spectral precision of auditory responses, influencing how sound features are encoded, integrated, and filtered. Alterations in this balance can disrupt auditory perception and have been implicated in disorders such as tinnitus and auditory hallucinations. Recent advances in computational neuroscience provide the means to estimate E/I ratio across the entire human cortex in vivo using biophysically informed large-scale models in combination with functional and structural MRI data (Zhang et al., 2024; DOI: 10.1073/pnas.2318641121). Such model-derived E/I estimates have been shown to be sensitive to pharmacological manipulation, cognitive performance, and aging. However, expression of these E/I ratios across human auditory cortex has remained understudied. Here, we address this question using resting state data from the Human Connectome Project and the methodology outlined in Zhang et al. (2024). Specifically, we aim to investigate the gradients of the E/I ratio within 14 auditory regions of interest using the Desikan-Killiany parcellation in relation to their neuroanatomical, functional, and structural properties.

Neuroanatomically, we confirm that E/I ratio across the entire cortex decreases significantly along the posterior-to-anterior axis and increases along the inferior-to-superior axis. However, within areas of auditory cortex, only the increase along the inferior-to-superior axis is significant. Functionally, similar to the cortex-wide expression of E/I ratios, the E/I ratio within the auditory cortex during resting state is well predicted by the position of a given region along the principal gradient of functional connectivity. Specifically, medial Heschl's gyrus, housing primary auditory cortex and situated at a relatively low level of functional hierarchy, exhibits highest E/I ratio. In contrast, middle temporal gyrus and temporal pole, positioned high in functional hierarchy, show the lowest E/I ratio of the auditory areas. Structurally, heterogeneity in E/I ratios at rest across auditory cortex scales linearly with the auditory regions' myelin content: the greater the myelination, the higher the E/I ratio.

These results demonstrate that within the human auditory cortex, the E/I ratio is heterogenous, and its gradient is shaped by the structural and functional architecture of the auditory regions. Future studies will be important in determining whether and how E/I ratios are modulated during auditory task performance or altered in auditory deficits.

ID: 217
Abstract
Keywords: Attention, Distraction, Oculomotor, EEG, Aging

Quantifying Sustained Attention and Susceptibility to Distraction using EEG and Eyetracking

Claudia Contadini-Wright, Maria Chait
Ear Institute, University College London, United Kingdom; c.wright.16@ucl.ac.uk

Listening in complex natural environments requires the concurrent ability to focus attention on target sounds (e.g.an announcement at the train station) whilst simultaneously ignoring task-irrelevant distractors (e.g the noise of the crowd). Uncovering the brain mechanisms which underly these abilities, how they relate to variability in performance outcomes, and how they may change with age, is crucial for understanding the challenges of real world listening.

To probe these processes, we designed an objective test of individual susceptibility to auditory distraction- validated using behavioural, electroencephalographic (EEG) and oculomotor (eyetracking) measures. We aim to quantify fluctuations in attentional state, thus elucidating the relationship between target enhancement and distractor suppression. As the ability to efficiently allocate attentional resources tends to decrease with age, we also investigate how the aging brain may attempt to compensate in chaotic listening environments.

Young (18-35) typical hearing listeners (N=50), and Older (60+) normal hearing adults (N=19; ongoing) engaged with a continuous performance task in the presence of frequent distractors. Subjects were tasked with identifying 'target' sounds (different across dimensions of frequency and duration) amongst a rich attended stream populated by pure-tone frequency deviants, length deviants and standard tones. As the task unfolded, EEG and oculomotor data were obtained simultaneously. Alpha and Theta activity, tonic arousal (via the Pupil Dilation Response), attentional allocation (via Microsaccade Rate) and ERPs were measured at points in time preceding either a correctly identified 'target' or a miss.

In younger adults, we found that microsaccade rate was lower in 3 second intervals preceding a hit, consistent with increased allocation of attentional resources. Preceding hits, evoked responses to sounds in the attended stream (standard tones, length deviants, frequency deviants) were significantly greater compared with evoked responses to the distractor stream. Conversely, responses to the distractor stream revealed greater activity in time periods preceding misses– implying that unsuccessful distractor suppression led to missed targets.

Alpha power was significantly higher preceding hits, consistent with an overall heightened state of arousal. Theta power was also significantly higher preceding hits, implying a successful, strategic and goal-directed application of attentional resources. Analysis of the older group is ongoing.

Together our results reveal an interplay of brain processes including concurrent distractor suppression and target enhancement that support successful listening. They also offer insights into how such processes may change as we age. Further analyses will focus on individual variability, connectivity, and correlations between these measures to determine whether distractor suppression and target enhancement are independent neural processes.

ID: 218
Abstract
Keywords: neural tracking, speech-in-noise, continuous speech processing, Auditory Steady State Responses, electroencephalography

The Effect of Masker Type on Neural Tracking of Continuous Speech in Noise

Maaike Smit[1], Ana B. Carbajal Chavez[1], Marlies Gillis[1], Maaike Vandermosten[1], Pol Ghesquière[2], Jan Wouters[1]
[1]ExpORL, Department of Neurosciences, KU Leuven, Leuven, Belgium; [2]Parenting and Special Education Research Unit, KU Leuven, Leuven, Belgium; maaike.smit@kuleuven.be

When processing speech, the brain time-locks to specific features of the incoming signal, a phenomenon known as 'neural tracking', the study of which offers insights into both acoustic and linguistic processing (Broderick et al., 2018). While the study of neural tracking of continuous speech has high ecological validity, paradigms eliciting responses to controlled auditory stimuli provide greater experimental control and robustness. Ideally, studies on spoken language processing incorporate both approaches. However, each requires extensive stimulation to elicit the target responses, resulting in lengthy experiment durations, which can be problematic for experiments with vulnerable populations, such as children or patients.

Speech rarely occurs in silence, and background noise affects both acoustic and semantic processing (Yasmin et al., 2023). Hence, studying speech-in-noise (SiN) processing is valuable in understanding how noise influences speech perception. Neural tracking of the speech envelope is modulated by the signal-to-noise ratio (SNR), whereby neural tracking is higher at low to moderate SNR levels and lower at higher SNR levels compared to speech in silence. Moreover, distinct noise types elicit different neural responses (Herrmann, 2025).

The SiN paradigm is widely used in the study of neural tracking of the speech envelope, as well as in the study of higher-order linguistic processing and (developmental) language disorders. Integrating controlled auditory stimuli as maskers within this paradigm could optimise the experimental design, thereby reducing experiment duration and making studies more feasible for vulnerable populations while preserving ecological validity.

In this study, we assess whether background noise can be leveraged to enable the simultaneous measurement of Auditory Steady State Responses (ASSRs) and neural tracking of continuous speech. We investigate how different noise types affect speech envelope tracking by using Sinusoidal Amplitude-Modulated (SAM) stimuli as masking noise in comparison to conventional Speech Weighted Noise (SWN). Our goal is to assess whether neural tracking of the speech envelope and ASSRs can be disentangled and reliably interpreted.

Participants are presented with speech at an SNR of 0 dB in three conditions: (1) speech in silence, (2) speech masked by a SAM stimulus, and (3) speech masked by SWN. By analysing neural responses across these conditions, we aim to determine whether measuring multiple aspects of neural processing in parallel gives robust and interpretable results. This approach has the potential to improve the efficiency of auditory neuroscience research, particularly for populations where long testing durations are impractical. Data collection is ongoing (aiming for N=25), with results to be presented at the upcoming conference.

ID: 219

Abstract

Keywords: nonhuman primates; auditory working memory; sensory cortex; prefrontal cortex; persistent activity

Persistent neuronal activity in auditory and prefrontal cortices is related to working memory in monkeys

Ying Huang, Michael Brosch

Leibniz Institute for Neurobiology, Germany; yhuang@lin-magdeburg.de

Persistent neuronal activity during working memory delay periods has been widely observed. However, its role in working memory has been questioned in recent years. In this study, we provide unequivocal evidence that persistent spiking activity in auditory and prefrontal cortices of monkeys is involved in holding information online during the delay period in auditory working memory tasks. Specifically, we recorded spiking activity of individual neurons from these brain regions while macaque monkeys performed a working memory task and a control task on sequences of two sounds S1 and S2. These two tasks differed in auditory working memory during the delay period between S1 and S2, but were similar with respect to other task-related mental processes such as stimulus or reward anticipation, motor preparation, and decision making. During the delay period, the spike rate of many neurons in auditory and prefrontal cortices differed between the two tasks. These differences persisted throughout the delay and were due to the difference in auditory working memory. These results support that persistent neuronal activity in sensory and prefrontal cortices underlies working memory.

ID: 220

Abstract

Keywords: nonhuman primates; working memory readout; sensory cortex; prefrontal cortex; match suppression

Match suppression in auditory and prefrontal cortices of monkeys during working memory readout

Ying Huang, Michael Brosch

Leibniz Institute for Neurobiology, Germany; yhuang@lin-magdeburg.de

While neuronal underpinnings of auditory working memory have been extensively studied during the storage of information, the readout phase—when the stored information is used to guide behavior to achieve goals—has received relatively little attention. The present study investigates spiking activity of individual neurons in auditory and prefrontal cortices of macaque monkeys while they utilize the information stored in working memory to determine whether the test tone is the same or different from the sample tone in a delayed match-to-sample task. During this readout phase, many neurons in the auditory and prefrontal cortices respond to the test tone differently when it matches or differs from the tone stored in working memory. Specifically, spike responses to the test tone are mainly suppressed when it matches the stored tone, which is the case in both auditory and prefrontal cortices. The match suppression in both brain regions reflects an active, task-related process rather than a passive one, because it occurs far less frequently when the same sound sequences are presented to the monkeys while they perform no tasks. These results indicate that both auditory and prefrontal cortices are involved in reading out the stored information so that it can be transformed into context-appropriate actions. These findings, together with our previous observation that these brain regions are also engaged in memory storage, suggest that distinct phases of the working memory process—namely storage and readout—can recruit the same neural circuit, spanning from low-order sensory cortices to high-order regions such as prefrontal cortex.

ID: 221

Abstract

Keywords: audition, auditory cortex, speech, invariant coding, intracranial EEG

Rapid and dynamic construction of acoustically invariant speech representations in the human auditory cortex

Dana Boebinger[1], David Skrill[1], Guoyang Liao[1], Kirill V. Nourski[2], Matthew A. Howard[2], Christopher M. Garcia[2], Thomas Wychowski[1], Webster Pilcher[1], Sam Norman-Haignere[1]

[1]University of Rochester; [2]The University of Iowa;
dana_boebinger@urmc.rochester.edu

Coding information in speech is computationally challenging in part because environmental factors such as reverberation, background sounds, and imperfect sound transmission can substantially alter the acoustics of speech. Much remains unknown about how the human auditory cortex represents speech information in a manner that is invariant to such acoustic variation, in part due to the coarse resolution of non-invasive neuroimaging methods such as functional MRI. Here, we measured neural responses to a diverse set of spoken sentences with and without common forms of acoustic variation, using spatiotemporally precise intracranial recordings from the auditory cortices of neurosurgical patients undergoing chronic invasive monitoring for medically intractable epilepsy. We developed a paradigm to measure the timing and strength of invariance relative to the onset of speech information. We found that the human auditory cortex constructs representations of speech that are strikingly invariant to reverberation, complex background sounds, and spectral variation. Invariance was particularly strong in speech-selective neural populations in non-primary regions of the middle portion of the superior temporal gyrus. While speech-selective electrodes tended to be located further away from primary auditory cortex than non-selective electrodes, the effect of speech-selectivity could not be explained by electrodes' position in the auditory hierarchy. In many neural populations, including speech-selective populations, invariance strength increased rapidly during the first 200 ms after the onset of speech information, indicating rapid and dynamic refinement of an initial acoustic representation. Importantly, the functional organization and time course of invariance could not be predicted by standard spectrotemporal receptive field (STRF) models. Across electrodes, we found that STRFs predicted a narrow range of invariance strengths and thus were unable to account for the variability in invariance strength we observed in different neural populations, including the effect of speech selectivity. Prior research has found that the auditory cortex gradually suppresses responses to stationary properties of sounds, but our results show that the auditory cortex rapidly constructs a highly invariant representation of speech even immediately after a change in acoustic variation. Collectively, these findings indicate that the human auditory cortex uses a rapid and dynamic process to compute a representation of speech-specific structure that is highly robust to many common forms of environmental acoustic variation.

ID: 222
Abstract
Keywords: auditory cortex, audition, invariance, speech, intracranial EEG

Representational disentangling of spoken words and talker identity in the human auditory cortex

Akhil Bandreddi[1], Dana Boebinger[1], David Skrill[1], Kirill V. Nourski[2], Matthew A. Howard[2], Christopher M. Garcia[2], Thomas Wychowski[2], Webster Pilcher[2], Sam Norman-Haignere[1]

[1]University of Rochester, United States of America; [2]The University of Iowa, United States of America; akhil_bandreddi@urmc.rochester.edu

Complex natural sounds such as speech contain many different sources of information, but recognizing these distinct information sources is computationally challenging because they are highly entangled in the acoustic waveform that reaches the ear. For example, variation in the acoustic attributes of different talkers makes it challenging to recognize the identity of a word, while variation in the acoustics of different words makes it challenging to recognize talker identity. How does the auditory system disentangle word identity from talker identity in auditory cortex? While disentanglement has been extensively studied in the visual cortex, less is known about how it is accomplished in the human auditory cortex. One hypothesis is that specific brain regions specialize for coding either word identity or talker identity. Alternatively, word identity and talker identity may be represented by distinct dimensions of the neural code at the population level instead of specific regions. Distinguishing between these two hypotheses has been challenging in part due to the coarse resolution of non-invasive neuroimaging methods such as fMRI. To address this question, we measured neural responses to a diverse set of 338 words spoken by 32 different talkers using spatiotemporally precise intracranial recordings from the auditory cortices of neurosurgical patients undergoing chronic invasive monitoring for medically intractable epilepsy. We developed a simple set of model-free experimental metrics for quantifying representational disentangling of word and talker identity, both within individual electrodes and brain regions, as well as across different dimensions of the neural population response. We observed individual electrodes in speech-selective regions of non-primary auditory cortex that show a representation of words that is partially robust to acoustic variation in talker identity, but no electrodes or brain regions showed a robust representation of talker identity. However, at the population level, we observed distinct dimensions of the neural response that nearly exclusively coded either words or talker identity. These results suggest that while there is partial specialization for talker-robust word identity in localized brain regions, robust disentangling is accomplished at the population level with distinct representations of words and talker identity mapped to distinct dimensions of the neural code for speech.

ID: 223

Abstract

Keywords: auditory verbal hallucinations (AVHs), voice perception, self-voice recognition, psychosis continuum, perceptual bias

Spatial and Identity Factors in Voice Perception: Implications for Understanding Auditory Verbal Hallucinations

Suong Tuyet Welp[1,2], Marc Schönwiesner[2]

[1]*Max Planck Institute for Human Cognitive and Brain Sciences, Germany;*
[2]*University of Leipzig;* welp@cbs.mpg.de

Auditory verbal hallucinations (AVHs) vary widely in their perceptual qualities, with differences in how voices are perceived spatially, identities attributed to them, and semantic content—suggesting both perceptual and cognitive mechanisms contribute to these experiences. This study investigates the hierarchical processing chain underlying voice perception by systematically manipulating ambiguous auditory stimuli (vocoded speech) across dimensions. We begin with fundamental spatial processing (internal vs external), progress to intermediate-level voice identity features, and culminate with higher-order semantic content analysis. This bottom-up approach allows us to isolate how each processing level contributes to voice perception and potentially to AVH experiences.

In Experiment 1, we tested 16 healthy, native German-speaking participants to examine how spatial presentation affects voice detection. Results showed that ambiguous stimuli were more readily identified as voices when presented from external locations versus internal locations. This finding may help explain why many AVHs are perceived as originating from external sources.

Experiment 2 investigated how voice identity and semantic content affect perception. Results revealed a clear hierarchical pattern: participants most readily perceived voices in stimuli based on their own voice, followed by strangers' voices saying the same words, then strangers' voices saying new words, and lowest voice perception for pseudo-words. This gradient demonstrates the combined influence of both voice identity (self vs. other) and semantic processing. These results align with working memory and voice identity models in schizophrenia literature.

Experiment 3 combined these factors by examining voice recognition accuracy across spatial locations. Participants demonstrated significantly better recognition of their own voice when presented internally compared to externally, consistent with the natural perception of one's voice during speech and suggesting an interaction between spatial and identity processing in voice perception.

Individual differences were analyzed through the Launay-Slade Hallucination Scale (LSHS) and Peters et al. Delusions Inventory (PDI). GLM regression analysis revealed that PDI scores significantly predicted both spatial biases and thresholds for voice perception at higher levels of vocoding. These findings suggest that even in non-clinical populations, cognitive patterns associated with delusional thinking may influence basic voice perception mechanisms.

Our results demonstrate how voice perception can be influenced by interactions between spatial processing, identity recognition, and individual cognitive predispositions. By identifying specific perceptual mechanisms potentially involved in AVHs, this research provides targets for investigating neural correlates of disordered voice perception and may inform therapeutic interventions for individuals experiencing AVHs.

ID: 224

Abstract

Keywords: MRI, diffusion, connectivity, language, structure-function

Structural predictors of receptive language selectivity in temporal cortex

Rebecca Michele Belisle, Yinuo Liu, Emily P. Stephen, Tyler K. Perrachione

Boston University, United States of America; rbelisle@bu.edu

The human brain's receptive language network consists of highly localized regions with strictly selective responses to language, distributed through the frontal and temporal lobes. However, the precise locations of these nodes vary considerably across individuals, and it remains unclear what factors lead language-specialized areas to arise in a particular location in any given brain. We hypothesize that unique patterns of structural features distinguish language-selective areas from adjacent non-language areas. Here, we test this hypothesis by training regularized logistic regression models (Lim & Hastie, 2013) to learn relationships between fMRI-localized receptive language regions in the temporal lobe and structural MRI-derived features (macroanatomical atlas regions, connectivity profiles, and microstructural measures). A receptive language localizer fMRI task (contrasting listening to intact vs. degraded speech) was used to identify language-selective functional regions of interest (fROIs) in individual (n=25) brains. These fROIs comprised the most language-selective vertices within four regions (aSTS, mid-STG, pSTS, and AngG) based on a probabilistic parcellation of the language network (n>800; Lipkin et al., 2022). Trained models were tested on left-out subjects, and performance was compared to two baseline models: one-subject-left-out group fMRI data and the aforementioned probability map. The structure-based models performed significantly better than chance (ROC-AUC>0.5), outperforming both baseline models for the majority of individuals. Notably, the structural features associated with each temporal parcel were quite distinct. For example, language-selectivity in aSTS was associated with greater connectivity with IFGoper., while selectivity in pSTS was associated with greater hippocampal connectivity, and selectivity in AngG was associated with lower connectivity with pallidum (a language non-responsive region; Wolna et al., 2025). The unique structural underpinnings of each fROI may reflect regionally distinct contributions to language processing. Interestingly, for all parcels, fROIs were more likely to be found within sulci, possibly paralleling the hypothesis that sulcal depth is related to functional specialization and cognitive task performance (Weiner, 2023). These findings form a first step towards understanding receptive language architectures in individual brains, which has implications for understanding post-stroke aphasia recovery, neurosurgical planning, and interindividual variation in language development and neurodevelopmental disorders. Furthermore, these findings can extend to comparative neuroanatomy by pinpointing architectural differences that support the uniquely human capacity of language. In sum, this data-driven identification of structural features underlying the language network supports the notion of shared architectural principles between individuals and has wide-ranging applications in language neuroscience.

ID: 225
Abstract
Keywords: fMRI, head motion, head stabilisation, imaging quality, motion artifacts

Prospective and retrospective evaluation of the efficacy of a new device to reduce head motion during auditory MRI: MR MinMo

Egor Levchenko[1,2], Raha Razin[1], Simon Richardson[1], Matthew Gardner[1], Hugo Chow-Wing-Bom[1], Kevin Hannon[4], James Murphy[4], Jeremy Skipper[1], Adam Tierney[2], Martina Callaghan[1], David Carmichael[3], Fred Dick[1]
[1]UCL, United Kingdom; [2]Birkbeck, University of London; [3]Kings College London; [4]A94 Technologies; e.levchenko@ucl.ac.uk

Head motion has widespread, pernicious effects on all MR imaging modalities, ones that only worsen with increased spatial resolution and higher magnetic field. Compared to visual cortex, the much smaller auditory cortex lies in brain regions already suffering from lower SNR and tSNR as well as substantial vascular artifact, making it particularly critical to reduce artifacts in order to increase sensitivity. Much scientific and technical effort has gone to developing computational algorithms to lessen motion corruption of acquired data. However, relatively little effort has been directed to preventing head motion occurring. This is despite the fact that this would circumvent position-dependent effects, such as variable sensitivity modulation of the receiver coil, shim changes, and spatially varying gradient performance, all of which are not corrected by retrospective motion artifact correction. Here, we introduce a new motion-minimization device (MR MinMo) and compare its performance to high-quality cushion sets commonly used in research and clinics worldwide. The MR MinMo device consists of a frame which conforms to the inner surface of the MRI head coil. It is fitted with a series of foam pads and inflatables ('modules'), mounted to most effectively and comfortably secure the participant's head. The frame has a hinged section which allows easy entry and egress; the inflatable modules are controlled by a custom manifold, hand-pumped by the MR operator to quickly and securely position the participant. The ear inflatables accommodate different earbuds for audio presentation, as well as being able to deliver sound pneumatically. In a pre-registered study, we used fast, whole-head EPI (TR=800ms) to estimate head motion with standard cushions and MR MinMo in 25 participants who made 10 different body movements typical of those observed during clinical and pediatric scanning (e.g. scratching the cheek, hip raise, leg cross, talking etc.). Across movement types, head motion was reduced by MR MinMo use. Using standard motion correction parameters as a metric, we found that head translations and rotations were on average smaller and more tightly distributed around a minimum when using the MR MinMo device, compared to the set of standard cushions. This was true for absolute displacement measures (reflecting shifting of the head over the entire run) as well as framewise displacements (volume by volume changes in position). Participants also rated the MR MinMo device highly on comfort at the beginning and end of the session. More generally, the device has been used for long (>>1 hour) functional and structural scanning sessions for several hundred child, adolescent adult healthy and clinical participants, with positive results. The ability to reduce motion sensitivity across lengthy imaging sessions will benefit auditory MRI, enabling deeper characterisation of function, microstructural organisation and the nuanced relationship between the two.

ID: 226

Abstract

Keywords: Plasticity Motion 7T High-Field-MRI fMRI Auditory Blind Vaso Layers

When visual areas become auditory areas: ultra-high-field fMRI (7T) reveals layer-dependent feedforward auditory motion processing in the hMT+/V5 of blind people.

Marco Barilari[1], Jacek Matuszewski[1], Remi Gau[2], Eléonore Giraudet[1], Marc Van Baelen[3], Renzo Huber[4], Olivier Collignon[1,5]

[1]Crossmodal Perception and Plasticity Lab, Institute of Neuroscience (IoNS) and Institute for Research in Psychological Sciences (IPSY), Université Catholique de Louvain (Belgium); [2]MIND, INRIA-Saclay, Palaiseau, Paris (France); [3]Department of Engineering Management, University of Antwerp, Antwerp (Belgium); [4]National Institutes of Mental Health, Bethesda, MD (USA); [5]HES-SO Valais-Walis, The Sense Innovation and Research Center, Lausanne and Sion, Switzerland; marco.barilari@uclouvain.be

The hMT+/V5, a region long thought to be specialized in visual motion, also responds to auditory motion in sighted and even more so in congenitally blind people. However, the mechanisms behind auditory motion information transfer to the "visual" motion area in sighted and congenitally blind individuals remain unknown. In this study, we characterized submillimeter activity in hMT+/V5 using ultra-high-field fMRI (7T) coupled with BOLD and VASO acquisition while sighted and congenitally blind processed visual and auditory motion information. More precisely, we inferred whether crossmodal motion-related information is processed through feedforward or feedback pathways by identifying circuitry from layer-dependent activity of hMT+/V5. Our study reveals that in sighted individuals, visual motion elicits a predominantly feedforward response within hMT+/V5, consistent with its role in the early visual motion processing. Interestingly, in congenitally blind individuals, we observe a similar feedforward response pattern for auditory motion, suggesting that in the absence of vision, auditory inputs engage hMT+/V5 in a manner analogous to visual motion processing for sighted individuals. These findings indicate that hMT+/V5 maintains a feedforward processing profile for motion despite crossmodal plasticity, with its sensory input shifting from vision to audition in the absence of early visual experience.

ID: 227

Abstract

Keywords: consciousness, human, iEEG, speaker-induced suppression, temporal response function

Intracranial Electrophysiology of Cortical Responses to Speech in Delirium

Emily R. Dappen[1], Alexander J. Billig[2], Ariane E. Rhone[1], Mitchell Steinschneider[1,3], Matthew I. Banks[4], Kirill V. Nourski[1]

[1]The University of Iowa, Iowa City, IA, USA; [2]The Ear Institute, University College London, London, UK; [3]Albert Einstein College of Medicine, New York, USA; [4]The University of Wisconsin-Madison, Madison, WI, USA; emily-dappen@uiowa.edu

Cortical speech processing involves multiple hierarchically organized regions. Higher-order areas involved in comprehension and monitoring of self-initiated speech are sensitive to changes in level of consciousness. Delirium is an acute and fluctuating disorder of consciousness characterized by inattention, disorganized thinking, confusion, emotional changes, hallucinations, and hypo- or hyperactive behaviors. Speech communication is impaired in delirium. This study examined cortical responses to self-generated and externally generated conversational speech during delirium.

Participants were adult neurosurgical patients undergoing intracranial electroencephalography (iEEG) monitoring for medically refractory epilepsy who were diagnosed with delirium following electrode implantation surgery or seizures. Clinical delirium assessments were administered twice daily on postoperative days 1-3 and following seizures. Audio recordings of the assessments and iEEG data were collected simultaneously. Suppression indices were calculated as the difference between the average broadband gamma (30-150 Hz) neural activity during the interviewer's speech and participant's own speech, divided by their sum. Suppression indices at each recording site were compared between delirium-negative and -positive conditions. Processing of conversational speech during the assessments was examined using a linear modeling approach (temporal response function, TRF). TRF output (model prediction accuracy) was compared between delirium-positive and -negative conditions for speech envelope, phoneme onset, word frequency, and semantic neighborhood density features.

Sites showing little to no difference in cortical activity between external and self-generated speech during delirium-negative condition often exhibited speaker-induced suppression or enhancement during delirium, reflecting a global imbalance in processing of self-generated speech. Auditory core cortex, middle frontal gyrus and sensorimotor cortex showed similar suppression indices in delirium-positive and -negative conditions. By contrast, other auditory areas within superior temporal cortex, as well as middle temporal and inferior frontal cortices showed greater suppression during delirium. TRF analysis revealed lower prediction accuracy in delirium for phone onset and semantic neighborhood density features but not for speech envelope.

Changes in suppression index and TRF model prediction accuracy are interpreted to reflect delirium-related impairments in cortical speech processing. Changes in suppression index during delirium may reflect impaired executive control and contribute to broader deficits in processing of self-related stimuli. Reduced TRF prediction accuracy suggests disruptions in speech processing beyond its basic acoustic attributes. This study expands our knowledge of the impact of delirium on cortical auditory processing. Future work will examine response latencies using a broader range of speech features.

ID: 228
Abstract
Keywords: speech perception, EEG, trust, temporal response functions

The Neural Consequences of Trust on Speech Perception

Jaimy Anne Hannah, Giovanni M. Di Liberto
Trinity College Dublin, The University of Dublin; hannahj@tcd.ie

Trust has a key role in human interaction supporting effective communication and cooperation. However, it is unclear how trust – or lack thereof – impacts how the human brain processes speech. While previous research has focused on how speech characteristics such as fluency or pitch contribute to the establishment of trust, the present study examines how trust in a communication partner impacts the neural processing of speech.

We used electroencephalography (EEG) to measure neural activity while participants listened to stories narrated by eight different speakers, after familiarising themselves with each speaker voice in a trust building phase. We hypothesised that trust would impact the neural processing of speech in several ways. First, we expected higher listening effort for untrustworthy speakers, as evidences by stronger neural tracking of the speech envelope and higher EEG alpha power. Second, we expected next-word predictions to be influenced by trust, with untrustworthy speakers increasing the prediction perplexity.

During the trust-building phase, participants played 20 trials of a trust-based investment game with each speaker, which were visually represented by “gnomes”, each with one specific voice and colour. Using gnomes prevented potential trust biases that would occur, for example, when using faces. On each trial, participants heard an investment rule (e.g., “I'll double your coins!”) and chose how many coins to invest from 1 to 10. Then, a certain number of coins was returned that may or may not match the promised value. Trustworthiness was manipulated by creating different lie distributions for each gnome, varying in frequency (20%, 50%, or 80% of the trials) and magnitude (returning a percentage of the promised value, ranging from 20% to 100%). After the trust-building phase, participants rated perceived trustworthiness for each gnome. They were then presented with monologues wherein each gnome speaks of a time they were accused of doing something wrong (e.g., cheating on a test). Participants were asked to rate how confident they were that the gnome was innocent.

The neural tracking of speech was estimated using multiple lagged linear regression on the story-listening EEG data. The regression included acoustic features, such as the speech envelope, and linguistic properties, such as next-word prediction surprisal, entropy, and perplexity estimated with an open-source large language model (LLM). This analysis produced temporal response functions (TRFs) for each feature, which were then analysed to measure the impact of trust on the neural processing of speech.

Differences in neural tracking were observed across trustworthiness conditions. Further, we found that LLMs were sensitive to trustworthiness as evidenced by differences in lexical prediction. These results provide insight into how the brain models trustworthiness in speech communication on a word-to-word basis.

ID: 229

Abstract

Keywords: auditory perception, repetition detection, auditory objects, sensorimotor synchronization

Auditory Object Formation in Temporally Complex Acoustic Scenes

Berfin Bastug[1,2,3], Yue Sun[3], Erich Schroger[1,2], David Poeppel[1,4]

[1]Max Planck School of Cognition, Germany; [2]Wilhelm-Wundt-Institute of Psychology, Leipzig University, Leipzig, Germany; [3]Ernst Strüngmann Institute for Neuroscience, Frankfurt am Main, Germany; [4]New York University, New York, USA; berfin.bastug@esi-frankfurt.de

The auditory system decomposes boundary-less sensory input into meaningful units through Auditory Scene Analysis (ASA) (Bregman, 1990). Repetition helps listeners segregate overlapping sounds and identify distinct auditory objects (McDermott et al., 2011). Previous studies suggest that repeated units in noisy or ambiguous contexts can eventually be perceived as stable auditory objects (Barczak et al., 2018; McDermott et al., 2011). However, the behavioral signature of this dynamic process remains largely unexplored. Here, we investigated this process by using “tone clouds”—randomly generated clusters of 50-ms tones lacking explicit boundary cues. We varied repetition strength by adjusting the ratio of repeated (target) to regenerated (distractor) tones. This created a continuum from fully repeated to continuous sequences, forming an auditory analogue of motion coherence tasks (Shadlen & Newsome, 1996). To perceive repetition, participants had to group repeated tones into stable auditory objects. This design allowed us to explore the minimal sensory evidence required for repetition perception in complex scenes. Participants completed two psychophysical experiments: repetition detection and sensorimotor synchronization (SMS). In the detection experiment, they judged whether a sound sequence contained repeated patterns. We also manipulated unit duration to better understand the nature of evidence required for stable object formation. In the SMS experiment, participants tapped in synchrony with the repeating pattern. Tapping behavior served as a real-time proxy for perceptual object formation and offered continuous behavioral insight into the build-up and stabilization of auditory objects. We show sigmoidal, quasi-categorical performance across repetition levels in both experiments. In detection, an interaction between unit duration and repetition strength is observed, with shorter durations showing earlier performance improvement. There is also a strategy shift. When repetition is unclear, participants appear to rely on fixed-duration timing; when repetition becomes clear, they switch to cycle-based timing. Interestingly, the interaction between unit duration and repetition disappears in decision time analysis. Once repetition is clearly perceivable, participants require the same cycle number (~ 4) to make a judgment, regardless of how long or information-rich the unit is. In the SMS, sigmoidal curves converge across unit durations, eliminating the interaction effect, and paralleling the detection time results. Within-trial progression analysis reveals a two-stage object formation process: when repetition is detectable, performance gradually builds up before reaching a saturation point, suggesting a categorical perceptual shift in strong repetition conditions, in which the additional evidence no longer enhances performance.

ID: 230
Abstract
Keywords: aging, change detection, auditory scene analysis, eye tracking, EEG

The Effect of Aging on Auditory Change Detection: EEG and Eye Tracking Investigations in humans

Mert Huviyetli, Maria Chait
Ear Institute, University College London, United Kingdom;
mert.huviyetli.21@ucl.ac.uk

The auditory system evolved as the brain's early warning system, continuously monitoring the acoustic environment and rapidly directing attention to new events. Because it is sensitive to a wide space inaccessible to other senses, the brain relies on auditory input to detect important changes in our surroundings—whether identifying a potential predator or prey in ancestral environments or responding to alarms and voices in modern settings. Here, we investigate how the brain's ability to monitor auditory scenes changes with age. Specifically, we focus on listeners' sensitivity to the emergence of new events and their capacity to use predictable structures in the environment to detect changes.

Previous behavioral studies (de Kerangal et al., 2021) have shown that while older adults generally exhibit impaired change detection, their sensitivity to predictability remains relatively intact. To explore how these behavioral patterns relate to underlying neural mechanisms, we conducted two separate experiments using EEG and eye tracking (N = 30 each) in passively listening participants.

Stimuli consisted of artificial auditory "scenes" composed of six concurrent tonal streams, simulating multiple simultaneous sound sources. In 50% of trials, a change event was introduced by adding a new source partway through the scene. Both the composition of the background scene and the identity of the added source varied across trials, ensuring that each scene was novel and that changes remained unpredictable.

While EEG analyses are ongoing, eye-tracking results suggest age-related impairments in auditory scene monitoring. Older listeners showed reduced sensitivity to predictability and altered interactions between attentional capture and phasic arousal in response to auditory scene changes. Notably, these effects were not associated with hearing thresholds, suggesting that these age-related listening difficulties reflect audiogram-independent aging processes that shape difficulties in listening.

ID: 231
Abstract
Keywords: natural sounds, categorization, population coding, representation, classifier

Auditory cortical manifold for natural soundscapes enables neurally aligned category decoding

Satyabrata Parida[1], Jereme C Wingert[1], Jonah Stickney[1], Sam V Norman-Haignere[2], Stephen V David[1]

[1]Oregon Health and Science University, United States of America; [2]University of Rochester, United States of America; davids@ohsu.edu

Auditory categorization is a crucial aspect of everyday communication, but we lack a clear understanding of the computations performed by the brain to define boundaries between groups of natural sounds. While modern deep learning-based sound classifiers perform as well as humans in categorizing sounds, they can rely on idiosyncratic acoustic cues that diverge from biologically relevant cues. We hypothesized that, since auditory cortical neurons show category selectivity, training a classifier with an auditory cortical front-end should more accurately reflect biological categorization. To test this, we trained an encoding model to predict single-unit spike train data from the ferret auditory cortex. To capture the immense heterogeneity in auditory cortical response properties and natural soundscapes, we used a multi-task architecture to train a single convolutional-neural-network-based encoding model across multiple neural recordings, each collected with different natural sound stimuli. The bottleneck layer of this comprehensive model yielded the “auditory cortical manifold,” which accounted for the spectrotemporal features driving responses of the entire cortical population. Representational similarity analysis showed that the manifold was consistent across animals and thus captured general cortical computations for natural sound processing. We pooled data across animals to train a single encoding model and used the resulting manifold to train a classifier. Additionally, we trained a spectrogram-based classifier (without neural grounding) and a neural classifier (using the spike count of auditory cortical neurons) on the same categorization task. Confusion patterns for the manifold-based classifier were better aligned with the neural classifier than with the spectrogram classifier, underscoring the value of using neural data to develop biologically grounded classifiers. These findings emphasize the potential of computational models that mimic sensory processing to achieve more biologically relevant performance.

ID: 232

Abstract

Keywords: Auditory Scene Analysis, Reverberation, Neural Speech Tracking, EEG, Selective Attention

Room for the Cocktail Party: Neural Tracking of Target and Distractor Speech in Reverberant Environments

Lars Hausfeld[1,2], Anna-Lena Krause[1], Sabina Wiley[1]

[1]Maastricht University, Dept. Cognitive Neuroscience, The Netherlands; [2]Maastricht Brain Imaging Centre (M-BIC), The Netherlands;
lars.hausfeld@maastrichtuniversity.nl

In environments with two concurrent speakers, listeners segregate a stream of interest (target) from a second stream (distractor). Our auditory system uses various cues – like spatial location or pitch – to support segregation and intelligibility of the target speech. While past studies investigated auditory processing in these environments using anechoic and often short speech, the neural processing underlying stream segregation under reverberant, naturalistic conditions remains unknown. In this study, we presented participants ($N = 18$) with 30-s audiobook excerpts of two concurrent speakers while their neural responses were measured with high-density EEG. Stimuli were manipulated to vary in reverberation (high vs. low), spatial location (co-located vs. separated), and pitch separation (small vs. large), using simulated binaural room impulse responses.

Behavioral results (intelligibility and perceived difficulty) showed a main effect of reverberation ($p < 10-8$) but not for location or pitch. To investigate cortical responses to the reverberant, continuous speech, we employed multivariate temporal response functions to predict EEG signals using envelope models of direct, or reverberant speech, or their combination (combined model). EEG prediction performance showed a main effect of reverberation ($p = .019$; high < low), in line with behavioral results, and of model ($p < 10-10$; reverberant < direct, combined).

For direct speech, interrogation of the combined model using cluster-based permutation tests ($p\_cluster < .05$) suggested an early stage of sound processing in fronto-central channels for target and distractor speech that were prolonged in high vs. low reverberation (low: 30-130ms; high 30-180ms). A subsequent left-lateralized attention effect (200-280ms) was observed for target vs. distractor speech in the low, but not high, reverberation condition, in line with behavioral results.

For reverberant speech, and low reverberation, we observed instantaneous processing of the target speech (-50-60ms) prior to the processing of the distractor (20-100ms; coinciding with the early processing stage for direct speech). Additional processing was found for reverberant target speech (200-370ms) but not the distractor. Differential processing of target vs. distractor speech (60-160ms) preceded that found for direct speech. These effects were not observed for reverberant speech in the high reverberation condition, indicative of increased difficulty and disrupted stream segregation.

Our findings suggest that in low reverberation, the auditory system can dissociate reverberant cues from target and distractor speech, potentially via a (subcortical) predictive mechanism for behaviorally relevant streams, supporting stream segregation and intelligibility. In high reverberation, however, reverberation might act as a general masker, reducing the ability of the auditory system to segregate streams effectively.

ID: 233
Abstract
Keywords: Cortical Gain Enhancement, Hearing loss, Development, EEG, Speech in Noise

Cortical Gain Enhancement in an Immature Central Auditory System

Talat Jabeen, Jami Fung, Bonnie Lau
Department of Otolaryngology, Head and Neck Surgery, University of Washington; tjabeen@uw.edu

Speech perception in noise is an important daily task for individuals across the lifespan. Children, in particular, rely on hearing in noisy classrooms to learn. While perceiving speech in noise is more challenging for all children due to central immaturities, it is even more challenging for hard of hearing children. Previous studies of cortical speech tracking in hard of hearing adults showed enhanced tracking of the speech envelope compared to the typically hearing adults, suggesting a central gain mechanism. To investigate whether enhanced cortical tracking of the speech envelope is observed in an immature central auditory system, we measured electroencephalography (EEG) responses in 7-to-18-year-old typically hearing and hard of hearing children. We simulated classroom listening conditions while recording EEG to an age-appropriate audiobook under 3 conditions: 1) Quiet – target alone at 0 degrees azimuth, 2) Co-located Noise – target and 4-talker babble at 0 degrees, 3) Separated Noise – target at 0 degrees, 4-talker babble at 90 degrees to the left and right. The EEG data were analyzed using the decoding model of the multivariate temporal response function (mTRF) toolbox in MATLAB. The preliminary results showed higher reconstruction accuracies for hard of hearing children in Co-located Noise but not Separated Noise. This finding suggests that while enhanced cortical tracking is observed under some noise conditions, further investigations are required to characterize cortical gain enhancement in the context of an immature central auditory system.

ID: 234

Abstract

Keywords: macaque, beat perception, sensorimotor synchronization, music, tapping

Monkeys have rhythm

Vani G. Rajendran[1,2], Juan Pablo Marquez[2], Luis Prado[2], Hugo Merchant[2]

[1]Institute of Cellular Physiology, National Autonomous University of Mexico (UNAM), Mexico City, Mexico; [2]Institute of Neurobiology, National Autonomous University of Mexico (UNAM) Campus Juriquilla, Querétaro, Mexico; vani.g.rajendran@gmail.com

Synchronizing movements to music is one of the hallmarks of human culture whose evolutionary and neurobiological origins remain unknown. The ability to synchronize movements requires 1) detecting a steady rhythmic pulse, or beat, out of a stream of complex sounds, 2) projecting this rhythmic pattern forward in time to predict future input, and 3) timing motor commands in anticipation of predicted future beats. Here, we demonstrate that the macaque is capable of synchronizing taps to a subjective beat in real music, and even spontaneously chooses to do so over alternative strategies. This contradicts the influential “vocal learning hypothesis” that musical beat synchronization is only possible in species with complex vocalizations such as humans and some songbirds. We propose an alternative view of musical beat perception and synchronization ability as a continuum onto which a wider range of species can be mapped depending on their ability to perform and coordinate the general abilities listed above through association with reward.

ID: 235
Abstract
Keywords: Auditory, Cortex, Deviant, Calcium imaging

Stimulus-specific adaptation (SSA) to pure tones in local networks of mouse auditory cortex

Dina Moshitch, Israel Nelken
The Hebrew University, Israel; dina.farkas@mail.huji.ac.il

Stimulus-specific adaptation (SSA) is the reduction in responses to a common stimulus that does not generalize, or only partially generalizes, to other, rare stimuli. In human studies, related brain potentials (such as Mismatch Negativity, MMN) are tested for a stronger property called ‘deviance sensitivity’. We test SSA for deviance sensitivity by comparing the responses to a deviant tone in an oddball sequence with two controls: the rare tone presented by itself (deviant-alone), and the rare tone presented as part of a sequence of many widely spaced frequencies (diverse broad). Deviance sensitivity requires that the responses in the deviant condition be larger than in the diverse broad condition.

Using fiber photometry and wide-field imaging of calcium signals in the mouse primary auditory cortex, we uncovered large and robust response components which show large responses to rare sounds, independent of their context (deviant, deviant-alone and diverse broad) and are often larger in the deviant than in the diverse broad condition.

To study the single-neuron activity that underlies these deviance-sensitive population responses, we used two-photon microscopy. We imaged neurons from upper layer 2/3 of the primary auditory cortex of mice. Neurons showed in general large responses to the three rare conditions (deviant, deviant-alone and the diverse broad) but formed distinct groups that seemed to prefer either deviant, deviant-alone, or diverse broad presentations of one frequency or the other. In consequence, population patterns distinguished all three rare conditions. Thus, the deviance of the deviant tones in an oddball sequence is reflected by specific population response patterns, in addition to the size of the responses they evoke.

ID: 236

Abstract

Keywords: sound processing and categorization, blindness, EEG, multivariate pattern analysis, Deep Neural Network

Temporal dynamics of sound representation in the brain of sighted and blind

Siddharth Talwar[1], Stefania Mattioni[2], Eléonore Giraudet[1], Roberta Calce[1], Francesca Barbero[1], Marco Barilari[1], Olivier Collignon[1,3]

[1]Institute for research in Psychology (IPSY) & Neuroscience (IoNS), Louvain Bionics, University of Louvain (UCLouvain), Louvain, Belgium; [2]Department of Experimental Psychology; UGent, Gent, Belgium; [3]HES-SO Valais-Wallis, School of Health Sciences; The Sense Innovation and Research Center, Lausanne and Sion, Switzerland; siddharth.talwar@uclouvain.be

Early blindness triggers a reorganization in brain networks supporting sound processing. How the temporal dynamics of auditory representations are impacted by visual deprivation remain however unknown. Using time-resolved multivariate pattern analyses applied to electroencephalic recordings (EEG), we reveal how blindness impacts the cascade of acoustic to semantic coding of natural sounds. Decoding analyses revealed a protracted enhancement of sound representation in congenitally blind (CB) in comparison to matched sighted controls (SC), predominantly localized to occipito-temporal regions. We then contrasted the ability of acoustic, categorical, behavioral, linguistic and sound-to-event deep neural network (DNN) representational models to predict neural responses to different sounds. We first show that models of early acoustic processes explained brain representations early in time with no differences between the two groups. Correlations between neural activity and some layers of DNN were enhanced in CB around ~200 ms after sound offset, likely representing modulations in intermediate acoustic processing that mostly mapped onto temporo-frontal regions. The categorical and behavioral representation of sounds were also enhanced in CB, peaking at 600 ms, and mostly mapped onto occipito-temporal regions. Finally, linguistic models peaked later in EB (around 1 sec) with no group differences observed. These results reveal that blindness triggers a selective reorganization of certain sound representations at specific timing of information processing, and mapping on to specific brain regions over time. This particular pattern of reorganization in brain networks coding for natural sounds in CB provides new insights on the reorganized dynamics of sound processing in absence of sight.

ID: 237
Abstract
Keywords: Predictive processing, mismatch negativity, local-global, entropy, audition

Investigating hierarchical auditory predictive processing with inharmonic sound sequences

Tomasz Domżalski[1], Alejandro Omar Blenkmann[2], Edyta Szurowska[3], Krzysztof Basiński[1]

[1]Auditory Neuroscience Laboratory, Department of Psychology, Medical University of Gdańsk, Gdańsk, Poland; [2]RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, Department of Psychology, University of Oslo, Oslo, Norway; [3]2nd Division of Radiology, Medical University of Gdańsk, Gdańsk, Poland; tomasz.domzalski@gumed.edu.pl

Predictive processing theory proposes that perception emerges from an interaction between top-down and bottom-up sensory input. At the core of this relationship is the internal model of the environment, which is updated with prediction errors (PE). A crucial component of this process is precision weighting, which describes the impact of PEs on model updating. Precision weighting is thought to depend on the entropy (amount of information, or simply predictability) of sensory inputs: lower - entropy environments enable more precise predictions, whereas higher - entropy environments lead to less precise, more uncertain predictions. The mismatch negativity (MMN) component of event - related potentials (ERP) is considered a neural marker of PE signaling. In this study, we investigate how entropy at different hierarchical levels of auditory processing affects MMN and P3 complex responses in a local - global paradigm. Participants are presented with three conditions: (a) low - entropy condition with sequences of harmonic complex tones where deviants differ only in fundamental frequency (f0), (b) a moderate - entropy condition with sequences of inharmonic complex tones, where sounds are jittered but maintain a fixed order, and (c) a high - entropy condition with sequences of inharmonic complex tones, where sounds are jittered and randomly ordered in a sequence. In all conditions, local and global deviants are introduced by omitting or altering the final element in a sequence. By varying sequence - level and stimulus - level entropy, we aim to dissociate stimulus - based uncertainty from sequence - based uncertainty and examine their effects on MMN and P3 complex. We hypothesize that MMN amplitude will be modulated by the predictability of both local and global rules, with larger PEs in low - entropy conditions due to higher precision weighting, and attenuated PEs in high - entropy conditions due to increased uncertainty. This study provides insight into how the brain integrates hierarchical uncertainty in predictive processing and contributes to better understanding of precision - weighted PE signaling in auditory cognition. The data collection is under way and the results will be presented at the ICAC meeting in Maastricht.

ID: 238

Abstract

Keywords: vocal perception, vocal production, electrocorticography, traveling wave, decision making

Global cortical dynamics during vocal communication in common marmosets (Callithrix jacchus)

Diya Yi[1,2], Misako Komatsu[3], Joji Tsunada[2]

[1]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China; [2]Chinese Institute for Brain Research, Beijing 102206, China; [3]Tokyo Institute of Technology, 226-8503 Tokyo, Japan; ydddelia@gmail.com

Vocal communication plays a vital role in maintaining group cohesion and social bonds among both humans and non-human primates. This process includes several cognitive steps: perceiving others' vocalizations, making decisions regarding vocal responses, and producing vocal motor outputs. Although cortical areas involved in these steps have been localized in humans and monkeys using electrophysiology and perturbation methods, the macro-scale cortical dynamics that link these cognitive processes remain unknown. To address this knowledge gap, we recorded whole-cortical activity using epidural electrocorticography (ECoG) in marmosets while subject monkeys vocally interacted with partners. We found that several cortical areas, including the parietal and auditory cortices, exhibited theta-band activity while listening to partners' calls. This sensory evoked activity was further modulated depending on whether the subjects engaged in vocal interactions with their partners, possibly representing decision-making processes and vocal motor preparation. Employing newly developed analytic techniques, Weakly Orthogonal Conjugate Contrast Analysis (WOCCA), we found that above cortical activity formed two distinct macro-scale cortical dynamics in the form of traveling waves: a rotational traveling wave corresponding to vocal preparation and a transitional traveling wave reflecting vocal decisions. Moreover, the power of transitional traveling wave was correlated with vocalization-induced suppressive activity in the high-gamma band, especially within the prefrontal and auditory cortices. Since vocalization-induced suppression depends on sensory prediction errors calculated based on expected auditory feedback of vocalizations, the transitional traveling wave may propagate specific decision- or acoustic-variables related to vocal production to local microcircuits. These results suggest that global traveling waves, potentially through coupling with local neural activity, orchestrate distributed representations of cognitive processes underlying vocal communication.

ID: 240

Abstract

Keywords: Musical behavior, development, rhythm perception, electroencephalography

Evidence for neural categorization of rhythm in human newborns

Francesca M. Barbero[1], Tomas Lenc[1,2], Alban Gallard[3], Nori Jacoby[4,5], Rainer Polak[6,7], Arthur Foulon[3], Sahar Moghimi[3], Sylvie Nozaradan[1,8]

[1]Institute of Neuroscience (IoNS), University of Louvain (UCLouvain), 1348 Louvain-la-Neuve, Belgium; [2]Basque Center on Cognition, Brain and Language (BCBL), Donostia-San Sebastian, Spain; [3]Groupe de Recherches sur l'Analyse Multimodale de la Fonction Cérébrale (GRAMFC, Inserm UMR1105), Université de Picardie, 80054 Amiens, France; [4]Computational Auditory Perception Group, Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt am Main, Germany; [5]Department of Psychology, Cornell University, Ithaca, NY 14853, United States; [6]RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Norway; [7]Department of Musicology, University of Oslo, Norway; [8]International Laboratory for Brain, Music and Sound Research (BRAMS), Montreal, Canada; francesca.barbero@uclouvain.be

Humans show an outstanding capacity to perceive, learn, and produce musical rhythms. These skills rely on mapping the infinite space of possible rhythmic sensory inputs onto a finite set of internal rhythm categories. What are the brain processes underlying rhythm categorization? One view is that rhythm categories stem from hard-wired neurobiological predispositions constraining internal representations of rhythmic inputs. However, a growing body of work suggests that rhythm categorization is plastic, that is, open to be shaped by experience over the course of life. To tease apart the relative contributions of neurobiological predispositions and experience in rhythm categorization, we measured neural responses to rhythm in healthy full-term human neonates, thus capitalizing on their minimal post-natal experience.

Scalp electroencephalography (EEG) was recorded from newborns (1 to 3 days old) while they were exposed to different acoustic sequences consisting of repeating patterns of two inter-onset intervals ranging from isochrony (1:1 interval ratio) to long-short patterns (2:1 ratio). In a second experiment, we separately recorded neural (EEG) and behavioral (sensorimotor synchronization) responses to the same rhythms in adult participants. The data were analyzed using a novel approach combining frequency-domain and representational similarity analyses (fRSA).

Preliminary results indicate significant rhythm categorization in neonates, with categories encompassing the 1:1 and 2:1 integer ratio rhythms, with a categorical structure remarkably similar to the neural and behavioral responses of adults. These findings suggest that internal representations of rhythm may be biased towards categorical structure by neurobiological mechanisms already in place at birth. Yet, these internal representations would be further built-upon and shaped by experience over the lifespan, leading to the diversity in music perception and behaviors observed worldwide. The current study thus constitutes a critical step toward clarifying the neural processes underlying rhythm perception. More generally, this research could contribute critical insights onto development and plasticity of perceptual processes specific to humans.

ID: 241

Abstract

Keywords: temporal prediction, pitch discrimination, temporal resolution, timing

Predicting when – how do temporal predictions benefit hearing?

Sophie Herbst[1,2,3]

[1]Cognitive Neuroimaging Unit, INSERM; [2]Neurospin, CEA Saclay; [3]Université Paris-Saclay; ksherbst@gmail.com

In a continuous stream of auditory signals, naturally characterized by temporal regularities, predicting the occurrence of future events is essential for adaptive listening. Extracting temporal regularities from auditory inputs allows the listener to form temporal predictions, that is to predict when a relevant event will occur, and focus capacity-limited resources on the predicted moments in time. Contrary to sensory predictions ('what'), temporal predictions ('when') cannot draw on a dedicated sensory system, but interact with the respective sensory analysis to enhance perceptual processing at the most relevant moments: temporal predictions thus guide the orienting of attention in time. Temporal predictions knowingly speed up motor and sensory responses, and enhance the sensitivity of the perceptual analysis in the auditory, visual, and tactile modality. Enhanced sensory responses at predicted moments distinguish temporal predictions from sensory predictions, which, in the strict sense of predictive coding surface as sensory suppression, rather than enhancement of neural signals when predictions are met.

Despite their relevance and ubiquity for the interaction with sensory environments, the cognitive and neural representations of temporal predictions derived from sensory inputs, and their functional consequences in particular for audition are not yet well understood. Here, I will present results from a set of studies that addressed the benefits from temporal predictions on the auditory sensory analysis in healthy human participants, across different tasks. While motor responses and pitch discrimination reliably benefit from temporal prediction in interval-based and rhythmic paradigms, other attributes like temporal resolution (gap detection) do not consistently improve. By combining these behavioral investigations with Bayesian modeling to formalize and quantify unfolding temporal predictions across trials, our results reveal specific synergies between temporal predictions and the sensory analysis of auditory events, while also highlighting aspects of the auditory analysis that appear less reliant on predicting when.

ID: 242

Abstract

Keywords: Microsaccades; auditory processing; auditory attention; pupillometry; eye movements

Oculomotor dynamics as a window into auditory selective attention

Xena Liu, Maria Chait

UCL Ear Institute, United Kingdom; xiner.liu.19@ucl.ac.uk

Background

Successful interaction with auditory information in our external world requires dynamic engagement of various cognitive processes, including arousal, attention, and memory. Recent research has established that such information can be uncovered from our eyes, through rapid orienting ocular responses which have been hypothesized to reflect arousal (phasic Pupil Dilation, PD), and attentional aspects (Microsaccades, MS) of auditory processing.

Methods

We simultaneously measured these two types of ocular data in several experiments (N=30 each), using an auditory working memory task with consecutive tone sequence triplets designed to temporally dissociate the processes of actively attending to behaviorally relevant targets and ignoring distractors. The main aim was to investigate whether previously observed MS dynamics are specific to active attention allocation, or generally sensitive to enhanced resource engagement including the process of distractor inhibition. Understanding such underpinnings of MS will provide deeper insights into auditory attention as well as potential guidance for research into challenges in effortful listening, many of which possibly arising from breakdowns in distractor suppression in real life auditory scenes.

Results

Distinct response patterns to individual tone sequences—reflecting fluctuations in instantaneous arousal and attentional engagement—were observed across PD, MS, and blink data. Each measure exhibited unique temporal characteristics, suggesting that they index different facets of attentive listening. Specifically, pupil dynamics (diameter and dilation rate) indicated heightened arousal during the distractor period, consistent with mechanisms of active suppression. In contrast, MS dynamics suggested that microsaccadic inhibition may be specifically linked to active attention, as this inhibition was notably reduced during the distractor period.

Discussion

The observed MS dynamics point toward potentially distinct neural mechanisms governing the modulation of selective attention during active listening versus active ignoring. Time-locked analyses, comparing correct and incorrect response trials, further revealed evidence for differential underlying processes across the various ocular measures during auditory processing.

Conclusion

Ocular metrics recorded during an auditory working memory task—particularly MS and related eye movement indices—demonstrate promise as tools for probing distinct mechanisms of selective auditory attention.

ID: 243

Abstract

Keywords: fMRI, Sound Delivery, Electronics, Equipment, High Fidelity

High fidelity, reliable, and comfortable earbuds for auditory fMRI at all field strengths: OMEMS (Oto Micro Electrical Mechanical Systems)

Oliver Josephs, Simon Richardson, Fred Dick

UCL, United Kingdom; o.josephs@ucl.ac.uk

Delivery of high quality audio for functional MRI using modern scanners is extremely challenging. Ideally, listeners would hear sound across the full range of human hearing (~30-10000+ Hz) at any amplitude level desired, with sounds free of harmonic distortion, ringing or aliasing, or additional spectral peaks. The sound would be delivered through comfortable, compliant and universally fitting earbuds that can be easily accommodated within snug head coils. The device would be entirely safe for any pulse sequence used, MR-invisible at the head, and audio performance would not be affected by RF transmit pulses or gradient switching. It would also be reliable, low-cost, and manufacturable and repairable with easy-to-source components, with a fully open source design.

Here, we present our ongoing work towards these lofty goals. The OMEMS device takes advantage of recent developments in MEMS (Micro-Electro-Mechanical System) speakers to assemble arrays of small drivers, nested in 3D-printed frames and protective, sound-damping housing. The speaker arrays for right and left ears are driven by standard electrical cables (thermally insulated from the body by a neoprene sleeve) connected to RF filters mounted on a standard penetration panel. Acoustic output is passed through short, relatively compliant pneumatic tubes to 3D printed acoustic couplers, which then connect to a compliant insert that fits snugly in the ear canal.

Electronics components (MEMS) are mounted on a custom-made, double-sided PCB, easily commercially manufacturable for a nominal price (~£5 per board for <10 units). The internal case for moutning the PCB and channelling the MEMS output is 3D-printed in SLS (Selective Laser Sintering) Nylon (PA1000), readily available from various sources. Ear connectors are also SLS-printed, and adapted to use high-quality, readily-available off-the-shelf eartips.

The OMEMS can deliver high-amplitude sound (up to safe stimulation levels) with < 1% harmonic distortion, relatively gentle spectral modulation even when uncorrected by the calculated inverse filter function, and reasonably full frequency response (up to 40-12000 Hz), with almost no pickup of gradient switching or RF energy (e.g., negligable change to RF reference voltage). Scanner noise attenuation from the earbud alone is >/= 17 dB overall, and can be augmented with additional padding and inflatables. Device components near the head impose minimal additional field inhomogeneity or MR signal, and safety and QA tests have shown no unsafe RF heating or MR artifacts.

The OMEMS device is currently in use with a Siemens 3T Prisma (32- and 64-channel coils) as well as the Siemens Terra 7T (with NOVA 32-channel coil).

ID: 245
Abstract
Keywords: predictive coding, statistical learning, EEG, computational modelling

Tracking the Statistical Structure of Rapidly Unfolding Sound Sequences - Evidence from EEG in Humans

Buse Adam, Maria Chait
Ear Institute, University College of London, London WC1X 8EE, UK; buse.adam.22@ucl.ac.uk

The brain is increasingly viewed as a statistical learner, as it monitors the structure of rapidly evolving sound sequences and makes predictions about the environment based on inferred statistics. An enduring question pertains to understanding which statistics the brain monitors automatically and how the representation of these statistics shapes the processing of new events.

A growing body of evidence suggests that the neural mechanisms involved in tracking auditory statistical regularities can be effectively studied through analyses of sustained M/EEG activity. These responses reliably vary with the predictability of sequential auditory inputs, offering a window into how the brain monitors and adapts to environmental statistics. In this study, we used the sustained neural response (across two experiments, N = 30 each) as a proxy for the brain's process of monitoring and adapting to auditory predictability. Our goal was to investigate which statistical features human listeners track and how recent auditory history influences neural responses.Predictions were guided by two computational models, PPM (Pearce, 2005) and DREX (Skerritt-Davis & Elhilali, 2018), representing different hypotheses concerning the brain's statistical tracking heuristics (Markov chain-based vs Bayesian ideal observer-based). Passively listening participants were exposed to rapid tone-pip sequences, containing occasional changes in the sequence statistics.

Experiment 1 involved stimulus sequences that transitioned between more and less predictable patterns, designed to elicit clear responses from both computational models—though with differing dynamics. In contrast, human EEG responses revealed selective sensitivity to certain statistical structures but not others, highlighting a limitation in the brain's capacity to automatically track all types of sequential patterns.

In Experiment 2, we investigated how the predictability of the pre-transition sequence influenced responses to identical tonal patterns. Stimuli began with either a predictable (regularly repeating) or a random—but otherwise acoustically matched—tone sequence, followed by a transition to a fixed pattern of 10 tones (REG10). We found that brain responses to the same REG10 pattern differed systematically depending on the preceding context. More predictable contexts led to later but less variable "transition" responses and distinct subsequent dynamics during the discovery of REG10.

Overall, the findings shed light on the intricate processes underlying the brain's ability to track statistics in rapidly evolving sound sequences, also highlighting the effects of previous context on the processing of changes in statistics.

ID: 246

Abstract

Keywords: Self-generation effects, perceived loudness, supra-threshold and near-threshold stimuli, auditory processing, N1

From suppression to enhancement: the effects of self-generation on sensory responses and perceived loudness depend on sound intensity

Carla Salgado[1,2], Stefanie Sturm[1,2], Luisa Maria Valencia Torres[1,2], Marc Via[1,2,3], Iria SanMiguel[1,2,3]

[1]Brainlab-Cognitive Neuroscience Research Group, Departament de Psicologia Clínica i Psicobiologia, Universitat de Barcelona, Barcelona, Spain; [2]Institut de Neurociències, Universitat de Barcelona, Barcelona, Spain; [3]Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain; luisavalencia2002@gmail.com

Self-generated (SG) sounds typically elicit attenuated perceptual and neural responses when presented at clearly audible (supra-threshold) intensities—a well-established phenomenon. Recent findings, however, suggest that this pattern reverses at near-threshold intensities, with SG sounds perceived as louder than externally generated ones. Despite this, few studies have jointly examined perceptual and neural effects of self-generation, and it remains unclear whether perceptual enhancements are mirrored in neural responses.

This study investigated how sound source (self vs. external) and sound intensity (near-threshold vs. supra-threshold) interactively influence auditory processing at both the perceptual and neural levels.

Forty-seven participants performed a two-alternative forced-choice intensity discrimination task on sound pairs. The first sound was either SG or externally generated, at fixed near- or supra-threshold intensity. The second sound was always externally generated, with variable intensity (±3 dB). Psychometric curves were fit to second-sound louder responses, and points of subjective equality (PSE) were extracted. Concurrently, EEG was recorded, and the amplitude of the N1 component elicited by the first sound was analyzed.

Sound source and intensity interactively modulated auditory processing. As expected, near-threshold sounds elicited weaker neural responses (lower N1 amplitudes and delayed N1 peaks) than supra-threshold sounds. However, intensity effects interacted with sound source effects. In general, the same interaction pattern was found on N1 amplitudes ($F(1,46) = 44.6$, $p < .001$) and PSE ($F(1,46) = 18.6$, $p < .001$): Sensory suppression was found for SG sounds at supra-threshold intensities, while sensory enhancement was found for SG sounds at near-threshold intensities. Remarkably, self-generation eliminated intensity-related differences in N1 amplitudes, effectively equalizing neural response amplitude across intensities.

This study expands upon prior behavioral research by integrating neurophysiological measures. The findings provide the first direct evidence that the reversal of self-generation effects at low intensities is reflected at both perceptual and neural levels, confirming a direct correspondence between these two measures. Moreover, the results highlight that action effects on sensory processing are not fixed but flexibly modulated by stimulus properties such as intensity. Results are discussed in the framework of predictive coding, with consideration of arousal-related contributions to sensorimotor integration processes.

ID: 247

Abstract

Keywords: music, rhythm perception, frequency tagging, fMRI

Similarities in processing simple and complex rhythms: behavioral and fMRI evidence

Ceren Battal[1,2], Tomas Lenc[1,5], Olivier Collignon[1,3,4], Sylvie Nozaradan[1]

[1]Institute of Neuroscience (IoNS), UCLouvain, Belgium; [2]Faculty of Psychology and Education, University of Coimbra, Portugal; [3]Institute of Psychology (IPSY), UCLouvain, Belgium; [4]The SENSE brain innovation and research centre, Lausanne, Switzerland; [5]Basque Center on Cognition, Brain and Language (BCBL), Donostia-San Sebastian, Spain; battal.ceren@gmail.com

When listening to musical rhythm, humans perceive and spontaneously move to a periodic pulse-like beat. Previous studies suggested the internal representation of a beat to be mainly driven by low-level properties of the acoustic stimulus, especially the prominence of a periodic pulse in the arrangement of sound onsets. This view has been motivated by the evidence for a functional sensorimotor network preferentially activated by simple rhythms containing prominent acoustic periodicities corresponding to the perceived beat, compared to complex rhythms that do not contain prominent acoustic periodicities and are assumed to weaken beat perception. However, prior studies often used short rhythmic sequences and behavioral tasks that might have confounded the data. Here, we challenge this view by showing that simple and complex rhythms played in longer ecologically-valid sequences lead to comparable performance in tapping along with the perceived beat. The same sequences were also played to the participants while performing an orthogonal task in separate functional MRI sessions. In line with tapping, both rhythm categories elicited similar activation of the auditory, premotor, and supplementary motor areas, despite the different amounts of temporal regularity in the stimulus. This lack of difference was not due to a limited sensitivity, as we obtained high signal-to-noise ratio when comparing all rhythmic sounds to silence. In addition, we observed robust pitch selectivity driven by modulating the pitch of the sound instead of rhythm category in a control experiment. Finally, the lack of difference between rhythm categories were not due to a lack of information content, as we could successfully differentiate simple and highly jittered/irregular rhythms in the auditory cortex. Together, our findings indicate that humans can build a robust internal representation of beat in the absence of unequivocal periodic cues, and that these temporally degraded stimuli equally activate the sensorimotor network.

ID: 248
Abstract
Keywords: Music Processing, Statistical Learning, Extracellular Electrophysiology, Multivariate Temporal Response Function

Statistical Learning of Melodies in the Auditory Cortex

John O'Doherty[1], Guilhem Marion[2], Hortense Gouyette[3], Giovanni M. Di Liberto[1,4], Yves Boubenec[3]
[1]School of Computer Science and Statistics, Trinity College, The University of Dublin, Ireland, ADAPT Centre, Ireland; [2]Department of Psychology, New York University, New York City, USA; [3]Laboratoire des Systèmes Perceptifs, Département d'Étude Cognitive, École Normale Supérieure, PSL, Paris, France; [4]Trinity College Institute of Neuroscience, Dublin, Ireland; odoherj1@tcd.ie

Processing melodies requires the brain to learn complex temporal and pitch structures. Humans learn culture-specific melodic structures from early ages through music presented by caregivers, such as nursey rhymes, allowing us to observe pitch discrimination and neural entrainment to the music beat in 7-month-old infants. This learning is facilitated by a phenomenon known as Statistical Learning (SL), whereby the brain learns regularities in the environment without conscious effort. The encoding of melodic statistics was measured and characterised in the human cortex. SL is a computational phenomenon broadly documented across mammals and birds, raising the intriguing possibility that learning musical statistics could extend to non-human animals. To explore this hypothesis, we investigated whether the auditory cortex of ferrets regularly exposed to music could similarly reflect an imprint of musical statistical structures.

The ferret is an ideal non-human subject for this investigation as its brain presents distinctions between primary and secondary areas in the auditory cortex that are similar to humans. Ferrets naïve to music were exposed to monophonic Bach chorales for three hours daily for 3 months, paired with positive contingencies. Then, neuronal responses to music were acquired with chronic arrays implanted over primary and secondary fields of the auditory cortex. Stimuli were exposed and left-out Bach pieces, and traditional Shanxi melodies, which has different statistical patterns. We then measured the neural encoding of music structure using multivariate lagged regression, producing Temporal Response Functions informing on how strongly music information was encoded in the neural signal, across different cortical sites and time-latencies. By controlling for the neural responses to sound, we could isolate the neural encoding of melodic expectations.

Neural signatures of melodic expectations emerged from the neural measurements of two ferrets, with stronger neural correlates for the exposed music. Timing statistics were more strongly encoded than pitch statistics, and Bach chorales are more strongly encoded than the Shanxi pieces. Our analyses were also sensitive to the learning that the ferrets experienced over the 10 experimental sessions, where their brain showed an increased encoding of the music structure of the Shanxi pieces by threefold.

Altogether, these results indicate that the ferret brain can learn music statistics with stronger sensitivity to timing statistics, in line with recent work in primates. Interestingly, learning effects were a result of passive exposure. In sum, this study demonstrates that the ferret brain can learn melodic statistics to a certain extent, offering precise insights into the spatial and temporal dynamics of that learning process. In run, we expect this research to open a new avenue for studying the learning of auditory development.

ID: 250
Abstract
Keywords: representations, temporal integration, recurrent neural networks

Investigating the computational principles underlying auditory cortical representations using recurrent networks

Lior Fox*[1,4], Arielle Rosinski*[1,4], Jennifer F. Linden[2,3], Maneesh Sahani[1]
[1]Gatsby Computational Neuroscience Unit, University College London, London, UK; [2]Ear Institute, University College London, London, UK; [3]Department of Neuroscience, Physiology & Pharmacology, University College London, London, UK; [4]equal contributions; lior.fox@ucl.ac.uk

What are the computational principles that underlie the formation of perceptual representations? Through development and learning, animals build internal representations of their sensory environments, which they use to guide and inform cognition and action. From an evolutionary standpoint, good representations are those that facilitate flexible and adaptive behaviour, but direct feedback about actions in the form of reinforcement or supervision is rare in nature. Thus, for animals at least, good representations may predominantly be shaped by unsupervised learning based on statistical regularities in sensory input—and indeed many experiments reveal changes in both behaviour and neural representations with passive exposure to altered sensory statistics, especially during early development.

Unsupervised learning is constrained both by representational or learning objective, and by architectural or circuit constraints. In the case of the auditory cortex, both of these principles are shaped by the fact that natural sounds are structured in temporal as well as spectral dimensions. Temporal structure makes it possible to consider predictive learning objectives (Singer et al., 2018) as well as reconstructive ones (Smith & Lewicki, 2006). But neural systems operate in real time, and so integration of temporal structure with memory beyond the time constant of a single neuron will depend on recurrent circuit mechanisms.

Here we address both constraints. First, we train recurrent models that process natural sounds sequentially, so that they are explicitly required to perform temporal integration. Second, we contrast prediction and reconstruction objectives. In a reconstruction task, the models are trained to process an input sound such that, for any moment in time, the past stimulus can be dynamically reconstructed from the current, integrated, hidden representation. This is in contrast with prediction-optimised models, in which the goal is predicting the future stimulus (from an integrated hidden representation).

Following training on a large database of natural sounds, we present the models with natural sounds that were previously used experimentally to characterize the population activity in the primary auditory cortex (Pennington and David, 2023). The hidden representations produced by the models, in response to a given stimulus, are then used as a set of linear predictors for the neural activity, recorded in response to the same stimulus, and we compare the different models by the quality of these predictions. Finally, we analyze model responses to a set of simple stimuli. We find that the response dynamics of reconstruction models exhibit longer timescales compared to prediction models, reflecting extended integration windows.

ID: 252
Abstract
Keywords: Selective attention, multisensory, audio-visual, electrophysiology, ferret

Auditory selective attention is impacted by the temporal coherence of a task-irrelevant visual stimulus in ferrets

Rebecca H C Norris, Zheyi Tang, Nadia Aghili, Jennifer K Bizley
UCL, United Kingdom; rhcnorris@gmail.com

Multisensory integration is a fundamental property of mammalian sensory systems, allowing us to combine information about objects in the world across senses. Audio-visual integration is especially relevant for understanding auditory processing, as visual information impacts auditory scene analysis, multi-modal object formation and speech processing.

We trained ferrets on an auditory selective attention go/no-go task, requiring them to respond to the presence of a brief timbre deviant embedded in a target stream while ignoring those in a simultaneous non-target stream. Both streams were independently amplitude modulated artificial vowels, with the target stream cued by beginning 1s earlier than the non-target. 25% of trials had no timbre deviants in either stream, requiring animals to hold for the full duration of the stimulus. Streams were 5 – 6.2s in duration, and target and non-target streams differed in both pitch and vowel identity (counterbalanced across trials). The non-target stream was attenuated relative to the target stream by 4-8 dB, with the relative level determined by the ferrets' performance, to avoid both floor and ceiling effects. Each trial had a luminance modulated visual stimulus in which luminance modulation was temporally coherent with amplitude modulation in either the target, non-target or neither stream. We selected visual stimulus properties as those showing the greatest impact on neural activity in auditory cortex in neural data recorded from ferrets presented passively with similar audio-visual stimuli. Importantly, the auditory feature coherent with the visual stimulus (amplitude) was orthogonal to the feature used for the detection task (timbre), and the visual stimulus provided no information that would assist the task.

Ferrets performed the task well, and response selectivity (d') was higher when the visual stimulus was coherent with the target stream compared to the non-target stream (0.94 – 1.52 vs 0.84 – 1.2, across 3 animals). The effect of visual coherence differed across animals, with d' changes in one animal driven by a higher hit rate in the target-coherent condition, while the two other animals additionally showed an increase in false alarms to non-target deviants when the visual stimulus was coherent with the non-target stream.

These results are in line with previous findings in a human version of this task, suggesting that temporal coherence results in the integration of visual and auditory information to create a cross-modal object. Ongoing work involves recording neural activity from auditory cortex during task performance to explore how audio-visual integration and selective attention interact to shape the representation of sound in auditory cortical neurons.

ID: 253

Abstract

Keywords: expectation suppression, predictive processing, laminar fMRI

Dissociable dynamic effects of expectation during statistical learning across cortical layers.

Hannah H McDermott[1,2,3], Mahdi Enan[2], Federico de Martino[2], Ryszard Auksztulewicz[1,2]

[1]Freie Universität Berlin, Germany; [2]Maastricht University, the Netherlands; [3]Berlin School of Mind and Brain, Germany; h.mcdermott@fu-berlin.de

The brain seemingly generates internal predictions, based e.g. on stimulus associations, to optimise behaviour. Predictive processing has been repeatedly demonstrated in non-invasive studies on human volunteers and in animal models. One commonly reported phenomenon is expectation suppression (ES) or the suppression of neural activity in response to expected stimuli. However, various mechanisms supporting ES have been suggested with conflicting evidence. Furthermore, most studies failed to demonstrate ES as a standalone phenomenon independent of repetition suppression. Our recent EEG study investigating the effects of predictions during associative learning shows that the effects of predictive processing are dynamic at both short and long time scales. In this high-field neuroimaging study, we test if these dissociable dynamics of expectation effects can be explained in the context of hierarchical learning mechanisms. Neuroimaging was performed using 7T fMRI, focusing on the primary/secondary auditory and prefrontal cortex (parcellated into layers) and the hippocampus (parcellated into subregions). During scanning, healthy volunteers (N=15) completed an associative learning task consisting of paired visual and auditory stimuli, whereby a “leading” image from a scene category is quickly followed by a “trailing” sound from a speech category (male vs. female voice) with 75% validity. Univariate analyses used an event-related general linear model to compare region- and layer-specific activity evoked by expected (valid) vs. unexpected (invalid) sounds. Multivariate analyses focussed on decoding accuracy in valid/invalid trials, and tested for decoding differences in the first half vs. second half of the experiment to quantify learning effects. Leave-one-run-out support vector machines (SVM) outline expectation effects on decoding accuracy, both within- and across- trials, between hierarchically lower vs. higher regions of the auditory cortical pathway, as well as between superficial vs. deep cortical layers. These results map the previously reported dynamic and opposing effects (sharpening vs. dampening) of stimulus prediction onto hierarchically lower vs. higher cortical regions, including contributions of specific cortical layers.

ID: 254
Abstract
Keywords: temporal processing, hearing loss, gap detection threshold, schizophrenia, 22q11.2 deletion syndrome

Distinct auditory temporal processing deficits contributed by hearing loss and genetic risk for schizophrenia in the Df1/+ mouse

Chen Lu[1], Zimo Li[2], Maneesh Sahani[3], Jennifer F. Linden[1,2]
[1]Ear Institute, University College London, London, UK; [2]Department of Neuroscience, Physiology, & Pharmacology, University College London, London, UK; [3]Gatsby Computational Neuroscience Unit, University College London, London, UK.; chen.lu.19@ucl.ac.uk

Central auditory processing dysfunction is common in schizophrenia patients. For example, duration thresholds for detection of a brief gap in noise are abnormally elevated in patients with psychosis (Iliadou et al., 2013). However, schizophrenia patients tend to have hearing impairment (Saperstein et al., 2023), and hearing impairment is itself associated with elevated gap-detection thresholds (Fitzgibbons and Wightman, 1982). Do temporal processing deficits associated with hearing impairment and genetic risk for schizophrenia arise from the same mechanisms?

Here, we used the Df1/+ mouse model of the 22q11.2 chromosomal microdeletion to address this question. In humans, the 22q11.2 deletion confers ~30% risk of developing schizophrenia and ~60% risk of hearing impairment, primarily from middle ear problems (Bassett and Chow, 2008; Verheij et al., 2018). The Df1/+ mouse has a homologous chromosomal microdeletion and recapitulates many features of the human deletion syndrome, including susceptibility to developmental middle ear problems (Fuchs et al., 2013; Lu & Linden, 2025).

We sought to disentangle the effects of early-onset hearing impairment and genetic risk for schizophrenia on auditory temporal processing by comparing cortical responses to gap-in-noise stimuli between Df1/+ mice with and without naturally occuring hearing impairment and WT mice with and without experimentally induced hearing impairment. To mimic middle ear problems affecting a subset of the Df1/+ mice, we performed ear surgery on WT mice at P11, removing the malleus bone or creating sham controls. Then, we recorded spiking activity of auditory cortical neurons using Neuropixels in awake, head-fixed adult mice passively listening to gap-in-noise stimuli with variable gap durations. We quantified cortical sensitivity to brief gaps in noise using robust methods we developed to identify onset and offset responses and deviations from sustained responses at both the single-neuron and neural population levels.

These analyses revealed distinct effects of hearing impairment and genetic risk for schizophrenia on different measures of temporal processing in the auditory cortex. In mice of either genotype with hearing impairment, both single-unit and population-level measures of neural sensitivity to brief gaps in noise revealed poorer temporal acuity, confirming results from human studies. Meanwhile, in Df1/+ mice with or without hearing impairment, fast-spiking units (putative parvalbumin-positive interneurons) exhibited increased sensitivity to noise onsets and offsets.

Our findings demonstrate that hearing impairment and the 22q11.2 deletion generate distinct effects on auditory cortical processing: hearing impairment broadly disrupts temporal acuity, while the Df1/+ deletion specifically alters fast-spiking interneuron dynamics. These results highlight separable mechanisms by which hearing impairment and genetic risk for schizophrenia may alter auditory temporal processing.

ID: 255
Abstract
Keywords: Stimulus-Specific Adaptation, auditory cortex, Inferior colliculus, Early responses, Late responses

Stimulus-Specific Adaptation in auditory cortex and Inferior colliculus

KAMINI SEHRAWAT, ISRAEL NELKEN
EDMOND AND LILY SAFRA CENTER FOR BRAIN SCIENCES, HEBREW UNIVERSITY OF JERUSALEM, ISRAEL; kamini.sehrawat@mail.huji.ac.il

Detecting changes in the environment is crucial for animal survival, and the mammalian brain is equipped with mechanisms to detect unexpected sound events. One example of such a detection in humans is mismatch negativity (MMN) which is an electrophysiological correlate of deviance detection, peaking around 150–200 ms after the onset of deviant sound. At the single-neuron level, Stimulus-specific adaptation (SSA) is a reduction in the response to a common stimulus that does not generalize or only partially generalizes to other rare stimuli. MMN and SSA are not identical, and a key distinction lies in their time courses: SSA overlaps the onset responses in auditory cortex and is therefore earlier than MMN.

We investigated SSA in the auditory cortex of mice using in vivo widefield calcium imaging. We found two distinct response components: an early component peaking within 50 ms of sound onset and a late component peaking at 100–150 ms that was not reported in single-neuron studies before. Notably, both components were present in the auditory cortex but the late component was absent in the inferior colliculus, indicating a cortical origin. These late responses reflect the averaged spiking activity of a local neuronal network that was deviance-sensitive, though their precise origin remains unknown. We compared these responses in two groups of mice: one had participated in a sound-preference task that included at least a few hours of exposure to a highly structured sound ensemble, while the other had never been exposed to sounds except for normal mouse housing environments. While substantial SSA was observed in both groups, the exposed mice showed significantly stronger adaptation and sensitivity to periodic contexts.

ID: 256
Abstract
Keywords: PAC, Tonotopy, Parcellation, fMRI, Connectomics

Investigating the sub-parcellation of the human primary auditory cortex using adaptational properties, cortical myelination and connectomics

Carl David Rushworth[1,2], Michael Asghar[3], Susan Francis[3], Katrin Krumbholz[1,2]
[1]Hearing Sciences, Mental Health and Clinical Neurosciences, School of Medicine, University of Nottingham, Nottingham NG7 2UH, UK; [2]NIHR Nottingham Biomedical Research Centre, Ropewalk House, 113 The Ropewalk, Nottingham NG1 5DU, UK.; [3]Sir Peter Mansfield Imaging Centre, School of Physics and Astronomy, University of Nottingham, Nottingham, UK; carl.rushworth@nottingham.ac.uk

Rationale:

New evidence from ultra-high field fMRI suggests that the human primary auditory cortex (PAC) is located obliquely along the long axis of Heschl's Gyrus (HG) and contains at least two sub-areas characterised by of tonotopic gradient reversals anteromedial (where more myelin has been detected) and posteromedial to the PAC (Besle et al., 2019).

The exact nature of these sub areas, however, is not yet clear, there are two key models of the PAC based on post-mortem studies which appear to coincide with these findings. Wallace et al. (2002), suggested the core area consists of two primary-like sub-areas located anterior and posterior to Heschl’s gyrus which correspond to the areas indicated by the gradient reversals, with the anterior area being more myelinated.

Alternatively, Morosan et al. (2001) suggested the anterior area itself is the core and the posterior area is a separate, secondary area rather than a sub area. They further suggest that the core is subdivided mediolaterally along the long axis of HG into areas they call TE 1.1, 1.2 and 1.3.

Similarly to other systems, there is a wealth of evidence to suggest that as the auditory pathway progresses, the adaptational properties of responses systematically change. One way in which this was operationalised was by Harms and Melcher (2002) who found that when presented with stimuli at increasing repetition rates, areas further along demonstrated more phasic responses, while areas earlier in the system remained tonic.

Our study intends to investigate the accuracy of Wallace and Morosan models by investigating differences in adaptational and structural properties between these sub-areas.

Method:

To do this, we have created a paradigm based on Harms and Melcher (2002), in which we use increasing repetitions of frequencies which will stimulate each tonotopic area separately. We will then investigate the differences in how phasic the responses between the areas are.

To further support this, we will also be collecting a quantitative T1 map (Sanchez-Panchuelo et al. 2021) which we will use as a marker for differences in myelination and use white-matter tractography based on diffusion-weighted imaging and resting-state functional connectivity to determine whether acoustic radiation projects more preferentially to one gradient or the other.

Hypotheses:

Although we are still collecting data for this study, we hypothesise that, should the Wallace model be correct: both areas will show similarly phasic response as the repetition rate increases, however, the anterior area will demonstrate greater myelination and preferential acoustic radiation compared to the posterior area. This would suggest two primary-like areas, one of which shows more primary traits.

Alternatively, should the Morosan model be correct: the anterior area will demonstrate a more tonic response as repetition rates increase, demonstrate more myelination and receive more acoustic radiation compare to the posterior area.

ID: 257

Abstract

Keywords: Cortical speech tracking, temporal response function, temporal processing

Role of sound onsets and offsets in cortical tracking of transient events during naturalistic speech listening

Adele Simon[1], Maria Chait[1], Jennifer F. Linden[1,2]

[1]Ear Institute, University College London, United Kingdom; [2]Department of Neuroscience, Physiology, and Pharmacology, University College London, London, UK; adele.simon@ucl.ac.uk

The amplitude envelope of speech is essential for effective comprehension. Research has shown that cortical activity in the theta-delta frequency bands tracks this amplitude envelope. The hypothesis underlying this tracking is that it reflects the cumulative response to transient events in the speech signal, such as sound onsets (the rapid increase in sound amplitude) and offsets (the rapid decrease in sound amplitude). While several studies have focused on sound onsets to investigate cortical auditory tracking, there has been little exploration of how sound offsets contribute to tracking speech signals.

In our study, we analysed two datasets from different laboratories, consisting of continuous EEG recordings from British (n=18) [2] and Danish (n=22) [3] participants listening to audiobooks. We used an onset/offset model developed to predict thalamic responses to temporally varying sounds in mice [1] to extract separately the onsets and offsets present in continuous speech. To assess the contribution of these transient sound events to the cortical response, we estimated Temporal Response Functions, which are weights for a linear model that maps a continuous stimulus to neural activity. We trained models using the representations of onsets, offsets, or a combination of both and compared the models' performances.

The performance of cortical speech tracking was quantified by correlating the measured EEG with the EEG predicted by our linear models. All models—those trained on onsets, offsets, and the combined onset-offset representation—demonstrated significant performance above chance levels (with $p<.001$). This finding indicates that both onsets and offsets contribute to the neural response when listening to speech. Notably, Wilcoxon tests used to compare prediction accuracy between different models across participants showed that models trained on onsets had higher prediction accuracy than those trained solely on offsets ($p<0.001$, Cohen's $d=0.69$), suggesting that amplitude increases have a more substantial impact than decreases. However, the combined model outperformed both the offset-only ($p<0.001$, Cohen's $d=0.85$) and onset-only ($p<0.001$, Cohen's $d=0.56$) models, underscoring the complementary role of sound offsets in cortical speech tracking.

Overall, our results suggest that cortical tracking of speech during continuous listening is driven by the neural responses to transient changes in sound amplitude, with sound onsets playing a dominant role while sound offsets also significantly contribute to the cortical response.

[1] Anderson LA & Linden JF (2016). J Neurosci 36(6), 1977-1995.

[2] Etard O & Reichenbach T (2022). 10.5281/zenodo.7086208

[3] Simon A, Bech S, Loquet G & Østergaard J (2022). 10.5281/zenodo.7500806

ID: 258
Abstract
Keywords: letter, number, integration, MVPA

Neural Activation Patterns of Letters and Numbers Overlap in the Temporo-Parieto-Occipital Junction

Zhiwei Chen[1], Francesco Gentile[1], Jan W. Kurzawski[1], Logan T. Dowdle[1], Dora Gozukara[2], Agustin Lage Castellanos[1], Milene Bonte[1]
[1]Maastricht Brain Imaging Center, Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands; [2]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Netherlands; zhiwei.chen@maastrichtuniversity.nl

Numeracy and literacy are fundamental cognitive skills that rely on associating visual symbols with their spoken representations. Prior research has identified the posterior temporal-parietal cortex as a key neural region for the cross-modal transformation of auditory and visual forms of alphanumeric symbols. However, the activation patterns underlying these transformations remain unclear. Here, we investigated the modality-independent cortical processing of spoken and written alphanumeric symbols. Twenty-one participants were presented with auditory or visual letters and numbers while performing a perception task in a slow-event-related 3T fMRI experiment. We found overlapping activation across auditory cortical regions for auditory letters/numbers and across ventral visual regions for visual letters/numbers. In particular, superior temporal cortical regions such as A5/A4/Parabelt (based on Glasser's HCP-MMP1 atlas) exhibited high reliability for auditory stimuli, whereas occipital and ventral temporal cortical regions such as V3/V4/PH demonstrated high reliability for visual stimuli. As univariate results showed no significant differences between BOLD responses to letters and numbers, we combined these two conditions for subsequent analyses. Interestingly, our results showed overlapping responses with similar amplitudes for both auditory and visual stimuli within the temporo-parieto-occipital junction (TPOJ). Furthermore, multivariate pattern analysis (MVPA) revealed that the TPOJ successfully differentiates between visual and auditory stimuli, achieving a decoding accuracy of 75% for the left hemisphere and 79% for the right hemisphere. Our findings reinforce the TPOJ's role in the cross-modal processing of symbolic representations. These results enhance our understanding of symbol processing in the brain and may have implications for developmental learning difficulties such as dyslexia, where cross-modal integration may form a challenge for acquiring reading fluency.

ID: 259
Abstract
Keywords: Auditory Cortex, Frontal Cortex, Auditory Streaming, Decision Making, Neuropixels

Representation, formation, and segregation of auditory streams from auditory to frontal cortices

Jeff Evan Boucher[1], Jules Lebert[1,2], Louise Dauron[1,3], Carla Griffiths[1], Jennifer Bizley[1]
[1]University College London, United Kingdom; [2]École Polytechnique Fédérale de Lausanne, Switzerland; [3]École normale supérieure, Paris; jeffreyevanboucher@gmail.com

The natural environment is filled with sounds that come from diverse sources and which have distinct behavioral meanings. In order to group the elements from physically distinct sound sources across time, our perceptual systems make use of temporal, spectral, and spatial cues in a process called Auditory Stream Segregation (Bregman 1990). While this process is well understood at the behavioral level through decades of research, the manner in which this process unfolds at the cellular level is still unclear. In particular, it is unknown in which brain regions and in what form streams are maintained, or how these factors depend on attentional focus or lack-of-focus. Additionally, it is unknown how the transition from a perception of "one, aggregate stream" into "two, split streams" may look in neocortex.

Our lab has been able to address these questions by making use of chronically implanted Neuropixels-2 probes in the frontal and auditory cortices of freely moving, behaving ferrets. This technology allows us to isolate single cells with high reliability and track them over multiple days. The ferrets were trained to recognize a target word presented at a variable point in time within a stream of randomly drawn non-target words. Ferrets performed this task in the presence or absence of speech shaped noise. The noise and speech were presented from different sides of space to encouraging stream splitting; however, they share a common temporal onset, promoting their initial integration. Moreover, the target stream's location was unknown at the start of each trial, requiring that the animal therefore initially attends to both locations. Using the temporal resolution of the Neuropixels probes, we can track the encoding of these streams in large populations of cells during these brief initial stages of stream-integration and stream-segregation, in both the auditory and the frontal cortex. As we are recording in frontal and auditory cortices simultaneously, we additionally aim to directly relate the activity patterns in these regions to one another. With this combination of complex behavioral paradigm and high channel count recording, we are able to address questions that have been inaccessible in the field for decades.

ID: 260
Abstract
Keywords: Neurofeedback, Speech Tracking, EEG, Temporal Response Function, Speech in Noise

Specific enhancement of neural responses to speech through neurofeedback

Ram K Pari[1,2], Marina Inyutina[1,3], Camille Dedies[2], Stefanie Enriquez-Geppert[4,5], Mathieu Marx[1,2,3], Christoph S Herrmann[6,7], Benedikt Zoefel[1,2]

[1]Centre de Recherche Cerveau et Cognition (CerCo), UMR 5549 CNRS , Toulouse, France; [2]University of Toulouse, France; [3]CHU Toulouse Purpan, Toulouse, France; [4]Department of Clinical and Developmental Neuropsychology, University of Groningen, The Netherlands; [5]Department of Biomedical Sciences of Cells & Systems, University of Groningen, The Netherlands; [6]Experimental Psychology Lab, Department of Psychology, European Medical School, Cluster of Excellence "Hearing4All", Carl von Ossietzky University, Oldenburg, Germany; [7]Research Center Neurosensory Science, Carl von Ossietzky University, Oldenburg, Germany; ram-kumar.pari@cnrs.fr

Deficits in speech comprehension are often caused by a reduced ability to suppress distracting speakers or noise. Successful attention to speech can be inferred from certain components of the "temporal response function" (TRF) that reliably differ between target and distracting speech. These components are also altered in ageing listeners and other populations that struggle to comprehend speech, making it an important marker for interventional approaches. The aim of the current study is to train participants to regulate these components through neurofeedback, possibly enhancing speech perception. We acquired Electroencephalography (EEG) data while participants listened to two audiobooks, presented simultaneously. We asked them to pay attention to one of them while ignoring the other and verified their attentional focus with comprehension questions. For audiobook snippets of ~22 seconds each, we extracted the N1 components (~110 ms) of the TRF, obtained in response to the derivative of the speech envelope for both target and distracting speech. These values were then expressed relative to a baseline condition (without neurofeedback) and preceding neurofeedback blocks and displayed to participants in the form of two visual bars (one for target, one for distracting speech). Participants were asked to regulate these bars to increase their neural response to target speech and minimising that to distracting speech. In preliminary data, we found an enhanced N1 component in the response to target speech that was only present during neurofeedback (N=15) but not during a subsequent baseline, nor in a control group (N=14; double-blind randomisation) that received a replay of neurofeedback values from other participants. No such change was observed in response to distracting speech. However, we found no clear effects of neurofeedback training on our measure of speech comprehension, possibly because this measure did not reliably capture cognitive functions that depend on the response component modulated with neurofeedback. Our preliminary results indicate that neural responses to speech can be regulated through neurofeedback, opening up new possibilities to manipulate these responses in populations where they are altered.

ID: 263

Abstract

Keywords: ribbon synapse, auditory coding. inner hair cells

Deficits in spike timing and transmission of auditory information in a mouse model lacking ribbon synapses

Wei Li, Najate Benamer, Saaid Safieddine, Jeremie Barral

wei.li@pasteur.fr, France; wei.li@pasteur.fr

In cochlear, inner hair cells (IHCs) specialized by the presence of ribbon synapses, its scaffolding protein-Ribeye tethers numerous synaptic vesicles anchored to the cell's active zone enables the rapid and sustained release of neurotransmitters essential for encoding sound with high temporal precision. Absence of Ribeye disrupts the synaptic ribbon and decreases the presynaptic density. At neonatal stages, compensatory development leads to an abnormal hearing threshold yet associated with an altered first peak in auditory brainstem responses (ABR). The exact role of ribbon synapses in sound encoding and the impact of their absence on later auditory pathway remain unknown.

This study conducted large-scale recordings from the cochlear nucleus (CN) and the inferior colliculus (IC) to assess neural differences between three genetic conditions: wild-type controls, neonatal knockouts (P2), and mature knockouts (P15).

Results demonstrated in both knockouts exhibits impaired synchronization with amplitude-modulated sounds in the CN and IC.

Dimensionality reduction analyses revealed increased intrinsic dimensionality in both regions of knockout models, also demonstrated a descrease of distinct coding regimes; in neonatal cKo, phase variability suggested more homogeneous in CN, while in the IC, decrease of sustained firing indicated a reliance on rate coding.

Although neuronal information could still be decoded using SVM or one-layer perceptron, the mechanisms varied between two recorded structures: decoding relied on spike timing in the CN and on firing rate in the IC. However, the information capacity was reduced in the CN but remained unaffected in the IC, suggesting compensatory mechanisms beyond the hair cell synapse, potentially extending along the auditory pathway. Dimension reduction analysis further indicated that intrinsic dimensionality was elevated in knockouts across both regions, highlighting alterations in neural coding properties due to the absence of ribbon synapses.

ID: 264

Abstract

Keywords: Mice, Discrimination, Harmonic sounds, Prediction

Behavioral discrimination between harmonic sounds in mice

Moran Aharoni[1,2], Lucas Drouet[1,2], Chloé Huetz[1,2], Jean-Marc Edeline[1,2]

[1]*University Paris-Saclay, France;* [2]*CNRS, France; moran.aharoni@mail.huji.ac.il*

Predictive processing is the default operation mode of our sensory systems for generating stable perception. In humans, these processes were prevalently studied using relatively non-invasive measures, such as EEG and MEG - and in rare cases, EcoG and electrophysiology. Studying prediction in animal models, specifically mice, allows for specific neural recordings and deeper understanding of the mechanisms behind the process. Experiment paradigms that induce prediction use sequences consisting of variable and discriminable stimuli (in our case, sounds). Currently, there are only a few studies looking at complex sound discrimination in mice. In order to develop a task that includes a large number of complex sounds, we ran a simple discrimination task between pairs of stimuli. Here, we study the discrimination as well as neural response to two pairs of sounds that include harmonics using a Go/No-Go protocol. The mice were head-fixed and received water reward when licking at the Go signal. The sounds comprising the pairs were 6 kHz, 8.5 kHz, 11 kHz and 13.5 kHz, and included the first two harmonics per sound. In separate experiments performed in anesthetized mice, we tested the responses to the single sounds comprising the two pairs: the responses to the four sounds were similar. Each sound pair had a frequency interval of 5kHz between the two sounds. We present the preliminary behavioral and neuronal results of the discrimination task between harmonic tone pairs.

ID: 265

Abstract

Keywords: spatial hearing, cross-modal perception, frequency-elevation mapping, music perception

Frequency-elevation mapping is driven by pitch rather than spectral content.

Zofia Anna Hołubowska[1,2], Marlen Herkenberg[1], Marc Schönwiesner[1,3]
[1]Institute of Biology, Faculty of Life Sciences, University of Leipzig, Leipzig, Germany; [2]Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; [3]Department of Psychology, University of Montreal, Montreal, Quebec, Canada; holubowska@cbs.mpg.de

Introduction & Aims

Frequency-elevation mapping (FEM) is an effect described for the first time almost a century ago (Pratt, 1930). In short, the location of high-pitched sounds is systematically overestimated, while lower sounds are underestimated. This effect seems to be driven by the adaptation to the statistics of the auditory scene, as well as the nature of the human auditory system, especially filtering properties of pinnae (Parise et al., 2014).

This study aims to determine whether the FEM is driven more by perceived pitch or spectral content (i.e., spectral centroid), and to explore how prior knowledge about sound affects spatial localization. We examine this locally - by assessing how the interval between preceding and target notes influences perceived location - and globally, by testing whether knowledge of an instrument's consistent spatial source across frequencies reduces the FEM effect.

Methods

We tested the Pratt effect using four timbre conditions: two musical instruments: (1) violin and (2) flute, and two artificial sounds: (3) harmonic tones with stable spectra, and (4) tones matching the viola's spectral profile. Conditions (1) - (4), and (2) - (3), were matched in spectral centroid, allowing grouping by either spectral content or sound type (musical vs. artificial). Participants ($n = 29$) heard tones at nine fundamental frequencies (200–1600 Hz) from five elevation angles (−25° to 25°) and indicated the perceived direction.

Results

We successfully replicated the Pratt effect, with fundamental frequency significantly affecting elevation perception across all conditions ($p < .001$). At low frequencies (~200 Hz), spectral centroid slightly increased perceived elevation (~1% variance explained), however this effect disappeared at higher frequencies. Local context mattered: the interval between notes significantly influenced elevation ($p = .018$), though this varied by condition. Globally, musical instrument sounds (violin, flute) showed weaker FEM slopes than artificial sounds, confirming reduced effects for musical conditions (both $p < .001$). Musical-based grouping better predicted elevation differences than centroid-based grouping (AIC: 102111 vs. 102135).

Discussion

Our findings confirm that FEM is a robust effect, largely independent of spectral centroid when pitch is clearly defined. While spectral features beyond fundamental frequency contribute only marginally, the effect appears to be more strongly shaped by context. Both local (e.g., preceding note frequency) and global (e.g., musical instruments as sound sources) factors influenced elevation perception, suggesting that FEM is not purely stimulus-driven. The notably reduced effect for musical instrument sounds implies that high-level knowledge - such as recognizing a stable

source capable of producing a range of pitches - can override low-level statistical regularities in spatial hearing. This highlights the role of top-down processing in shaping auditory perception.

ID: 266

Abstract

Keywords: Deep learning, Synergy, Redundancy, Natural sounds, Information theory

Auditory redundant and synergistic representations between brain and artificial neural networks revealed by information-theoretic decomposition

Christian Ferreyra[1,2], Marie Plegat[1], Giorgio Marinato[1], Maria de Araújo Vitória[3], Michele Esposito[3], Elia Formisano[3], Thierry Artières[2], Bruno L. Giordano[1]

[1]Institut de Neurosciences de la Timone, CNRS, Aix-Marseille Univ., Marseille, France.; [2]Laboratoire d'Informatique et des Systèmes, Aix Marseille Université, UMR 7020 CNRS, Marseille, France.; [3]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands.; christian.ferreyra@etu.univ-amu.fr

Artificial neural networks (ANNs) have become increasingly useful for modeling the cerebral representation of natural sounds (Giordano et al., 2024). The nature of their representational alignment with dynamic brain activity remains however comparatively unexplored. Here, we introduced an information-theoretic approach to analyze the temporal interplay between human auditory brain representations and candidate ANNs of the acoustic-to-semantic transformation. We combined magnetoencephalography (MEG) recordings from participants listening to natural sounds (2s long), with two sound-processing ANNs with categorical (CatDNN) and continuous (SemDNN) semantic outputs (Esposito et al., 2024), and analyzed the alignment of cerebral responses with model representations for two optimized stimulus sets.

We computed representational dissimilarity matrices (RDMs) from sensor-level MEG signals and from ANN activations. Then, we used partial information decomposition (PID; Williams and Beer, 2010) to quantify the time-varying mutual information between brain and model RDMs, and, importantly, the redundant (shared) and synergistic (complementary) interactions between model-based representations in time-resolved MEG RDMs. Critically, we tested these alignments using stimuli sets optimized to either maximize or minimize models agreement.

For low-agreement stimuli, the intermediate SemDNN layer aligned more strongly with brain activity than for CatDNN throughout the sound duration. When focusing on SemDNN, PID modelling revealed strong redundancy and synergy, the latter suggesting sustained temporal integration of continuous intermediate semantic representations. Redundant interactions were strongest during the first second of the cerebral response and between nearby time points, while synergistic information maintained long-range interactions throughout the sound duration. For the high-agreement stimulus set, both redundant and synergistic interactions showed similar patterns but with weaker effect sizes, and with redundant effects being notably concentrated at sound onset.

In contrast, CatDNN showed limited redundant and synergistic alignment under low-agreement conditions. In the high-agreement regime, CatDNN showed similar redundant interactions than SemDNN, and slightly weaker synergistic interactions, indicating potentially weaker temporal integration for CatDNN representation even in the absence of similarly strong representations at each latency of the brain response.

Overall, we observed a clear representation and temporal interaction of intermediate continuous semantic attributes of natural sounds (SemDNN) in dynamic cerebral responses. Synergistic interactions suggest that the temporal integration of these

semantic representations could afford a more accurate readout of the auditory environment. Our work highlights the potential of PID as a powerful framework for disentangling shared and complementary contributions of neural and model representations.

ID: 267
Abstract
Keywords: fast fMRI, auditory processing, high resolution, HRF modelling

Characterising the auditory hemodynamic response function in the era of fast fMRI

Letitia Maria Schneider[1], Isma Zulfiqar[1], Yaël Balbastre[1], Lori Holt[2], Martina Callaghan[1], Tim Schmidt[3], Zoltan Nagy[3], Frederic Dick[1]
[1]University College London, United Kingdom; [2]University of Texas at Austin, United States of America; [3]University of Zurich, Switzerland; l.schneider@ucl.ac.uk

We can now use fMRI to image the whole brain with higher spatial (</=2mm voxel size) and temporal (</= 1000 ms) resolution and sensitivity. Consequently, it may be possible to gain additional information about underlying neural responses by estimating and applying individuals' hemodynamic response functions (HRFs) at regional or voxelwise levels. Recent vision research indicates visual HRFs to short-duration stimuli are faster and narrower than predicted by canonical HRF models, with systematic associations between HRFs and the anterior-posterior axis of V1. Given the importance of temporal information in the auditory system at various timescales, as well as the intricate vasculature in auditory cortex, we conducted a set of studies to characterize voxelwise HRFs to short, naturalistic auditory stimuli, where we examined reproducibility across sessions, within and between participants, and across auditory-relevant regions. We also explored whether increased temporal resolution and signal-to-noise ratio (SNR) sampling can provide additional insights about auditory processing. In the first study, five healthy adults completed two identical sessions of passive listening to short environmental sounds at 3T (6 runs each, TR=1s, 2mm isotropic voxels). In a second study, four participants from the first study plus three additional ones completed four sessions (6 runs each) optimized for a new technique that achieves high temporal resolution (~80ms) with high SNR (volume TR=1s, 2.5mm isotropic voxels). Voxelwise HRFs were computed after performing between-session nonlinear brain alignment with a novel nonlinear warping method. Responses were estimated across the entire brain using gentle (2mm) cortical-surface-based smoothing of EPI timeseries after masking out large blood vessels. We identified various responses, including complex shapes in event-related averages that canonical HRF models (SPM standard plus derivatives) could not capture, which were highly consistent across sessions per participant. We developed six HRF models (using 1, 2, and 3 gamma functions) to capture the diverse hemodynamic shapes observed, fitting models to voxelwise responses averaged across stimulus repetitions. Even data from just two runs allowed HRF models (using both simulated and measured data) to explain 60% variance of the true HRF (excluding one more-variable participant). Auditory regions along the temporal lobe displayed highly consistent responses over time and exhibited patterns in time-to-peak and response widths replicable across participants. Specifically, faster and sharper responses were localized more medially, whereas slower and wider responses occurred more laterally. We also tested this pattern for responses sampled at a higher temporal resolution in the larger dataset. Finally, we assessed whether sampling at higher versus lower temporal resolution could reliably capture additional information about the response, such as transient changes.

ID: 268

Abstract

Keywords: Auditory expectations, Magnetoencephalography (MEG), Perceptual inference, hearing in noise

How predictions and update signals shape our auditory experience

Yamil Vidal[1], Jorie van Haren[2], Federico De Martino[2], Floris P. de Lange[1,3]

[1]Donders Centre for Cognitive Neuroimaging, Netherlands, The; [2]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands; [3]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands;

yamil.vidal@donders.ru.nl

Perception is often cast as an inferential process in which the brain combines noisy sensory input with predictions generated by internal models. Different algorithmic incarnations have been proposed to explain how this combination could take place. In them, along a hierarchy of generative models, feedback connections carry predictions from higher level to lower level areas, and feedforward connections carry either prediction errors, or likelihoods that can be used as an update signal.

It has been proposed that the contents of our perceptual experience are determined by the current hypothesis about the state of the world that has the highest posterior probability. Following the above, the content of our experience should be determined by the interactions of predictions (priors) and update signals, but there is currently no agreement regarding how these signals are combined.

We have designed a task in which expected and unexpected auditory stimuli were presented against a noisy background, rendering their detection challenging, and therefore bringing to the foreground the inferential nature of perception. By acquiring magnetoencephalography (MEG) and behavioural data during the execution of this auditory task, we can relate the amplitude of update signals to the detectability of tones.

We tested the hypothesis that the detection of expected tones should be associated with lower amplitude evoked fields, because sensory evidence matches with the prediction at hand. Instead, the detection of unexpected tones would only be possible when they elicit and update signal to successfully override the mistaken prediction. Importantly, evoked fields amplitude should be lower when an unexpected stimulus is not detected, suggesting that the internal model failed to update.

Behavioural results show that participants have higher perceptual sensitivity (higher d') and a more liberal response bias (lower criterion) when detecting expected tones. They are also faster at detecting tones that they expect.

In order to isolate update signals in the MEG, we subtracted the evoked fields elicited by expected tones from the ones corresponding to unexpected tones, independently for detected and missed tones. We found that this update signal has higher amplitude when tones are detected ,compared to missed tones, over a cluster at around 350ms after stimulus onset and with a topography suggesting a source in the auditory cortex.

We also found that the main effect of detection appears earlier than the main effect of expectation. This suggests that contrary to what has been proposed, the detection of tones is not determined by the “winning hypothesis” after model update. Instead, tones are first detected, and the update of the internal model happens afterwards.

With this study, we hope to clarify how predictions and update signals interact to determine the conscious perception of sounds, addressing a central question in the study of auditory perception.

ID: 269

Abstract

Keywords: auditory corticothalamic circuitry, auditory attention, computational model, thalamic reticular nucleus

A neural mass model of the auditory corticothalamic circuitry

Matthijs Kusters[1], Mario Senden[2], Michelle Moerel[1,2]

[1]Maastricht Centre for Systems Biology and Bioinformatics (MaCSBio), Faculty of Science and Engineering, Maastricht University, the Netherlands; [2]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, the Netherlands; mjh.kusters@student.maastrichtuniversity.nl

Previous studies have suggested that the circuitry between the auditory thalamus and cortex may play a crucial role in auditory attention[1–3]. However, it is not know if, and under what constraints, this circuitry can adaptively optimize neural tuning for the task at hand. Here we build a computational model of the auditory corticothalamic circuitry, including the medial geniculate body (MGB), primary auditory cortex (PAC), to address this question. Each brain region was modeled using Wilson-Cowan neural mass equations[4], with 98 coupled excitatory and inhibitory populations representing the tonotopic axis in MGB and PAC, and 98 inhibitory populations for the TRN. The model architecture included bidirectional connections between MGB and PAC, and unidirectional connections from PAC to TRN and from TRN to MGB. We optimized the model parameters, controlling intra- and interregional dynamics (the gain and spatial spread of connections), through a sensitivity analysis via 40,000 simulations sampling a wide range of parameter combinations. Of the 40,000 simulations, 2,710 parameter sets maintained the model in an active transient state. From these, we selected the parameter set that produced firing rates and tuning widths best matching empirical electrophysiological data from animal studies[5,6]. We then further optimized the parameters through dense sampling of the local parameter space. The final model produced robust and appropriately tuned responses across all three regions. Specifically, in response to pure tones the model produced firing rates and tuning widths of 25.6 Hz (0.146 oct) in MGB, 45.8 Hz (0.123 oct) in PAC, and 17.0 Hz (0.380 oct) in TRN. Our next step will be to incorporate frequency-specific attention in the model, to test whether the circuitry can account for results of previous studies of auditory attention[7–9].

1. Homma & Bajo. Front Neurosci 15, (2021).
2. Guo, Clause, Barth-Maron & Polley. Neuron 95, 180-194.e5 (2017).
3. Williamson & Polley. Elife 8, (2019).
4. Zulfiqar, Moerel & Formisano. Front Comput Neurosci 13, (2020).
5. Bartlett, Sadagopan & Wang. J Neurophysiol 106, 849–859 (2011).
6. Simm, De Ribaupierre, De Ribaupierre & Rouiller. J Neurophysiol 63, 1010–1021 (1990).
7. Fritz, Elhilali, David & Shamma. Curr Opin Neurobiol 17, 437–455 (2007).
8. Riecke et al. Neuroimage 174, 274–287 (2018).
9. Lage-Castellanos et al.. Cerebral Cortex 33, 5395 (2023).

ID: 270

Abstract

Keywords: adolescence, speech perception in noise, neurocognitive development, magnetic resonance spectroscopy, diffusion weighted imaging

Neuroplasticity in speech-in-noise processing across pubertal stages: A longitudinal study of adolescent cognitive and auditory development

Joseph Francois Johnson[1], Marta Puertollano[1], Francesca Cavicchiolo[1], Luna Prud'homme[1], Antonin Rovai[2,3], Vincent Wens[2,3], Simon Dobri[4], Xavier De Tiège[2,3], Axelle Calcus[1]

[1]Université Libre de Bruxelles, Center for Research in Cognition & Neurosciences, Laboratoire Cognition Langage et Développement, Belgium; [2]Université Libre de Bruxelles, ULB Neuroscience Institute, Laboratoire de Neuroanatomie et de Neuroimagerie translationnelles, Belgium; [3]Université Libre de Bruxelles, Hôpital Universitaire de Bruxelles, CUB Hôpital Erasme, Department of Translational Neuroimaging, Belgium; [4]Simon Fraser University, Canada; jfj.research@gmail.com

Adolescence represents an extended period of neuroplasticity that promotes experience-driven adaptations across brain regions and networks supporting the cognitive and behavioural capacities required in dynamic social environments. During this developmental phase, adolescents increasingly engage with larger peer groups, which is supported by their ability to track and discriminate voice signals in noisy backgrounds or in the presence of competing speech sounds. Such skill relies on a network of functionally and structurally connected regions including prefrontal cognitive control, parietal cross-modal association, and the fine-tuning of temporal auditory voice perception. Age-related brain changes during adolescence are well documented. However, defining how and when these changes occur remains challenging due to emergent pubertal hormones. As part of the SensationaHL pubertal development cohort collection, we aim to investigate the neuroplastic mechanisms and myelination trajectories linked to pubertal stages and hormonal markers, focusing on brain regions serving complex cognitive and auditory skills. A two-years longitudinal cohort of 236 participants will be recruited, assessing pubertal stage transition surrounding puberty onset and offset. Pubertal hormone levels and stage assessments will be collected at three timepoints (baseline and one-year intervals), in addition to cognitive and auditory speech-perception tasks. Multimodal neuroimaging including single-voxel spectroscopy, diffusion imaging, and electroencephalography will be used at baseline and after two years to evaluate neurotransmitters, structural and functional connectivity and thus infer upon the mechanisms of neuroplasticity involved. We hypothesize that myelination trajectories will correlate with pubertal stage progression, extending past puberty-offset and into late adolescence. We also predict an association between improvement in complex auditory skills, executive function, and puberty-associated gamma- aminobutryic acid (GABA) variations, a key modulator of neuroplasticity.

ID: 271
Abstract
Keywords: cognitive control, multiple demand, speech in noise, fMRI, functional connectivity

Probing Connectivity among Auditory Cortex Subdivisions and the Multiple Demand Network During Listening to Clear and Degraded Speech

Ali Tafakkor[1,2], Madison Tutton[1,2,3], Aysha Motala[4], Björn Herrmann[5,6], Ingrid Johnsrude[1,2,3]
[1]Western Institute for Neuroscience, Western University, London, Canada.; [2]Center for Brain and Mind, Western University, London, Canada.; [3]Department of Psychology, Western University, London, Canada.; [4]Department of Psychology, University of Stirling, Scotland, UK; [5]Rotman Research Institute, Baycrest Academy for Research and Education, Toronto, Canada.; [6]Department of Psychology, University of Toronto, Toronto, Canada; atafakko@uwo.ca

Speech listening often happens with other sounds in the background, requiring the auditory system to draw on higher-order cognitive resources for successful comprehension. Domain-general ”multiple demand” (MD) regions (Duncan, 2010), including the anterior insula (Ins) and anterior cingulate cortex (ACC), are known to be engaged when speech becomes challenging to comprehend (Ritz et al., 2022). Frontal regions are densely interconnected anatomically with auditory and other temporal-lobe regions (Davis & Johnsrude, 2007): speech perception is likely mediated at least in part by frontotemporal circuits. Here we examine how coupling between frontal and temporal regions shifts as background noise level increases. Conventionally, “primary AC” is often treated as a single functional entity, yet anatomical investigations in humans and other primates indicates multiple primary regions surrounded by a belt of secondary regions, and tertiary and quaternary regions in temporal and frontal cortex (Hackett et al, 2011). A cytoarchitectonic analysis of primary AC in humans (Morosan et al., 2001) reveals three distinct subregions within the core auditory field (“core”: Te1.1, Te1.0, Te1.2). Te1.0 is the most primary-like region (Hackett et al., 2001), Te1.2 is a more anterolateral region, perhaps corresponding to area R or RT in the nonhuman primate (Moerel et al., 2014), whereas area Te1.1 may be a medial belt region (Hackett et al 2001). These areas may exhibit different functional coupling with MD regions during perception of clear and degraded speech. Here, we use fMRI to examine functional connections among MD regions, as well as pairwise coupling among MD regions and the three “primary” auditory regions, while participants listened to engaging stories, presented both clearly and at 4 signal-to-noise ratios (SNR). Functional MRI data were collected from 44 normal-hearing adults (aged 18–35) who listened to three 10–14-minute stories (from The Moth podcast) in the presence of multi-talker babble noise. The SNR varied across five levels (clear, +14, +9, +4, −1 dB), creating a range of listening demands. After realignment and normalization to a common template (MNI152), a factorial general linear model tested for effects across the five noise conditions, including nuisance regressors (e.g., motion and acoustic envelope). Participants’ functional data was then analyzed at the group level using voxel wise contrasts and cluster-level corrections.

Our preliminary analysis indicates that activity in Te1.1, but not 1.0 or 1.2, increases as SNR increases. Conversely, activity in Ins and ACC increases as SNR decreases. Further analysis will examine functional coupling among these five regions, and how it changes with noise level. By examining coupling between frontal cognitive control regions and these AC subdivisions, we hope to clarify how sensory and cognitive-

control circuits work together to maintain intelligibility in challenging auditory environments.

ID: 272
Abstract
Keywords: adaptation, computational modelling, electrophysiology, magnetoencephalography, auditory sensory memory

Lifetimes of sensory memory traces in the auditory cortex of gerbils and humans: an interdisciplinary investigation

Nina Härtwich[1], Matthias Deliano[1], Eike Budinger[1,2], Frank W. Ohl[1,2,3], Asim H. Dar[1], Ewelina Tomana[4], Cezary Sielużycki[4], Reinhard König[1], Patrick J.C. May[5,1]

[1]Leibniz Institute for Neurobiology, Magdeburg, Germany; [2]Center for Behavioral Brain Sciences, Magdeburg, Germany; [3]Otto von Guericke University, Magdeburg, Germany; [4]Wrocław University of Science and Technology, Wrocław, Poland; [5]Lancaster University, Lancaster, United Kingdom; nina.haertwich@lin-magdeburg.de

While memory traces in the auditory cortex (AC) can be observed behaviourally, the underlying neural mechanisms are less obvious. One indirect observation of sensory memory at the neural level is repetition suppression or adaptation, a decrease in response amplitude when a stimulus is presented repeatedly. This indicates that memory traces of past stimuli affect responses to incoming stimuli.

A key mechanism of adaptation is short-term synaptic depression (STSD), but factors that modulate the lifetime of the reflected memory trace remain unclear. Therefore, we carried out an interdisciplinary investigation combining computational modelling of the AC with in-vivo recordings, covering two species (Mongolian gerbil, human) and two levels of observation (electrophysiology, magnetoencephalography (MEG)). We also created 1) a bootstrap-based analysis pipeline to produce robust estimates of in-vivo adaptation lifetimes at the single-subject level and 2) an optimisation pipeline to improve a computational AC model, based on an evolutionary algorithm.

A unique feature of our modelling approach is that it reflects the species-specific core-belt-parabelt network pattern that characterises the mammalian AC. The response dynamics of our new gerbil AC model suggest that, rather than being a simple reflection of STSD lifetime, the adaptation lifetime for individual neural populations is modulated by the interplay between STSD dynamics and network connectivity patterns. A matching prediction resulted from our optimised human AC model: when network connection strengths were altered during the optimisation process, adaptation lifetimes changed, even though model parameters affecting STSD dynamics were fixed. Furthermore, the gerbil AC model predicted that network interactions cause changes in adaptation lifetime as a function of audio-frequency. This was partially confirmed via intracranial recordings from the primary auditory field of six gerbils. The data set allowed us to produce the first detailed reports of adaptation lifetimes for individual neural populations in the AC. In four out of six animals, lifetimes were audio-frequency specific. Moreover, they were significantly shorter than the lifetimes deduced from MEG recordings in the ACs of 14 human subjects. The analysis of this MEG data also revealed a, so far unreported, hemispheric difference in adaptation, with longer lifetimes in the left than the right AC.

Our findings suggest that adaptation in the AC is strongly modulated by network effects. Variations in adaptation lifetime resulting from network interactions might functionally contribute to the temporal binding of sounds. The adaptation lifetime difference observed across species suggests that time windows for temporal binding match the shorter time scales of gerbil vocalisations and the longer time scales of

human speech. Similarly, hemisphere-specific adaptation lifetime could reflect larger integration time windows in the left than the right AC.

ID: 273

Abstract

Keywords: human, speech, eeg, perception, production

Rapid neurodynamic adjustments to cortical speech encoding influence speech production

Lori L. Holt[1], Fernando Llanos[1], Timothy K. Murphy[2], Yunan Charles Wu[3], Craig A. Thorburn[1], Lin N. Zhou[1], Frederic Dick[4], Nazbanou Nozari[5]

[1]University of Texas at Austin, United States of America; [2]University of Wisconsin-Madison, United States of America; [3]Carnegie Mellon University, United States of America; [4]University College London, United Kingdom; [5]Indiana University, United States of America; lori.holt@austin.utexas.edu

Perceptual processing of acoustic speech affects speech motor control. Rapid, directionally predictable adjustments to speech production occur when auditory feedback is altered, and these are well-explained by neurobiologically plausible models. Speech produced by others also shapes production. Yet current models, designed around self-generated feedback, don't account for these cross-talker perception-production effects.

To examine this, we leverage the strong correlation between voice onset time (VOT) and fundamental frequency (F0) in American English. Normally, both dimensions inform /b/-/p/ categorization. But when we reverse the correlation to simulate a subtle accent, F0 no longer signals category identity. Using EEG, we find that exposure to such distributional regularities shifts cortical responses across multiple timescales. This influence emerges at early latencies (N1) linked to sound feature encoding in auditory cortex. Strikingly, cortical differentiation between speech stimuli with large F0 differences (80 Hz) vanishes in the context of an accent. Correspondingly, mid-latency responses (MMN) across passive listening show that these acoustically distinct stimuli are more poorly differentiated in accented versus canonical contexts. Thus, statistical learning drives cortical encoding to down-weight dimensions deviating from long-term norms.

This rapid auditory plasticity also alters speech production. The perceptual down-weighting of F0 is mirrored in speakers' own utterances: in accented contexts, listeners use F0 less distinctively when producing /b/-/p/. This suggests that auditory cortical representations guiding motor control are malleable, not fixed, targets as assumed by many models. Notably, this effect doesn't require prior overt production; the altered use of F0 appears in the first post-exposure utterance. Contrary to some influential models, perceptual processing can impact production without the production system's prior engagement.

These results call for new ways of looking at a fundamental question: how perception affects action. To this end, we have begun to examine the nature of the perceptual representations that influence production. Further results suggest general auditory contributions: nonspeech tones carrying no articulatory-phonetic information shift speech perception, with a concomitant influence on speech production. These spectrally contrastive perceptual effects, common across speech and nonspeech, impact speech motor control. Thus, production is not sequestered from general auditory perceptual processing. Seemingly 'low level' perceptual processes that diminish distinctiveness in cortical encoding or sharpen contrast across successive sounds impact speech motor control.

ID: 274
Abstract
Keywords: Auditory perception, discrimination, generalization

Auditory Discrimination and Generalization in Mice Across Different Behavioral Paradigms

Fabian Ottolin[1,2], Irene Onorato[1], Livia de Hoz[1,3]
[1]Charité Medical University, Neuroscience Research Center, Berlin, Germany; [2]International Graduate Program Medical Neurosciences, Charité Medical University, Berlin, Germany; [3]Bernstein Center for Computational Neuroscience, Berlin, Germany; fabian.ottolin@charite.de

To better understand auditory perception in mice, we compared sound discrimination and generalization across two different behavioral paradigms: a two-alternative forced-choice (2AFC) task and a delayed match-to-sample (DMTS) task. Both tasks were conducted on a floating platform, which allows head-fixed mice to interact with behavior ports by moving the platform with their feet. This enables concurrent acute Neuropixels recordings in a more naturalistic setting.

In the 2AFC task, mice were trained to discriminate between two pure tones one octave apart in frequency. Each tone was assigned a fixed behavioral response (left/right port), this was maintained for all experimental sessions. After the mice learned to discriminate the sounds, generalization to different tones with frequencies above, between, and below the two pure tones was assessed to determine the mice's perceptual framework. Each trial was freely initiated by nose-poking into the center port. During the discrimination phase, depending on the stimulus, approaching the correct port was rewarded. For the generalization phase, all new stimuli were rewarded independently of the response to not bias the psychometric curve.

In the DMTS task, mice learned to indicate whether they perceived the current sound as being the same or different as the previously presented one, thus it also includes some aspects of auditory discrimination. Whereas in the 2AFC task the correct response depended on a fixed value assigned to each sound (left/right), in the DMTS task the same sound was rewarded or not based on its relation to the previous sound. Therefore, the learned association with reward was more dependent on context. The task involved discrimination between multiple sounds (vs. two sounds in the 2AFC task). Moreover, the DMTS task also comprises working memory aspects, as it is necessary to keep representations of the previous sound active until hearing the next stimulus. Mice were rewarded if they correctly indicated that the previously presented sound differed from the currently presented one by interacting with the response port. In trials where the current and previous sound were the same, mice were not rewarded and could initiate the next trial by poking into the sample port. Poking into the response port when current and previous sounds were the same, or the sample port in trials where the sounds differed, was punished by a time-out period.

Since the cognitive demands of the two tasks differ, we hope to gain insights into how this affects sensory processing and sound representations. In the DMTS, we also presented complex sounds, such as naturalistic stimuli, in addition to pure tones. We plan to assess whether generalization and discrimination vary between sound exemplars and categories. Combining different behavioral paradigms to investigate sensory coding and perceptual learning may provide novel insights into how task demands moderate neural representations of auditory percepts.

ID: 275

Abstract

Keywords: Voices, Voice Familiarity, Perceptual Organization, Attention Decoding, EEG

Neural correlates of selective attention to familiar voices.

Bruno Mesquita[1], Tobias Reichenbach[2], Ingrid Johnsrude[1]

[1]Brain and Mind, Western University; [2]Friedrich-Alexander-Universität Erlangen-Nürnberg; bmesquit@uwo.ca

Several studies demonstrate how familiar voices are more intelligible when masking sound is present (Johnsrude et al., 2013; Holmes & Johnsrude, 2020). However, higher intelligibility may arise due to any number of underlying mechanisms. Interestingly, some studies have also shown that the familiarity effect offered a benefit not only for attending a familiar target, but also in ignoring a familiar masker to selectively attend to a target. These have been demonstrated both in as an improvement in intelligibility (Johnsrude et al., 2013), as well as in lower working memory load when a distractor voice is familiar (Kreitewolf et al., 2019; Fischer et al., in prep), and it is hypothesized that this effect is only present for highly familiar voices.

That an intelligibility benefit can arise when the task-irrelevant, ignored, voice is familiar may hint at the familiarity effect enhancing perceptual organization of the auditory scene (i.e., greater perceptual segregation of voices), rather than using a familiar-voice template to more accurately identify words.

To date, the work in this area has been primarily behavioral, and it is unclear whether the brain dynamically tracks familiar voices differently from unfamiliar ones. Here, we used EEG and decoding methods to examine the strength of neural coupling to two-talker voice mixtures, when either the target or masking talker were naturally familiar to the participant.

We recruited 8 pairs of participants, all long-term (>1 year) heterosexual couples to visit our laboratory and record their voices as they read different stories. Participants were then brought back to the laboratory to listen to recordings made by their partner and other participants while we recorded EEG data. The experimental session was divided into five blocks. During the first two blocks, participants listened to two talkers reading different stories at the same time. In one of the 20-min blocks (FC), one talker was familiar (F) and the other (C) was unfamiliar, and from a member of the opposite sex to the F. Every 2 minutes, the participant was instructed to switch their attention from F to C or vice versa. During the other block (NC), they heard the same C voice speaking a different story, now paired with another unfamiliar talker of the same sex as F (N). In the last three blocks, participants listened to a single talker (F, C, and N) at a time, reciting different stories. Participants answered comprehension questions about the stories in between each 2-minute trial.

Preliminary analysis shows decoding of selective attention to the attended stories by comparing the reconstruction of envelope modulation of the attended and ignored talkers in each trial. Limited comparisons between familiarity conditions indicate no difference between decoding accuracy towards attended voices, but higher decoding performance towards familiar voices when ignored. Further analyses will probe the differences between familiar and unfamiliar voices.

ID: 276

Abstract

Keywords: multi-talker, audio-tactile, multisensory, magnetoencephalography, speech-in-noise

Vibrotactile enhancement of attended speech cortical tracking in the supratemporal auditory cortex in a cocktail-party scenario

Sabina Rautu[1,2], Mathieu Bourguignon[1,2], Vincent Wens[1,2,3], Julie Bertels[1,2], Xavier De Tiège[1,2,3]

[1]Universite libre de Bruxelles, Belgium; [2]ULB Neuroscience Institute, Bruxelles, Belgium; [3]Hôpital Erasme, Bruxelles, Belgium; ioana-sabina.rautu@ulb.be

Understanding speech in environments with multiple competing speakers—i.e., cocktail-party scenarios—is challenging. In such complex auditory scenes, providing listeners with rhythmic, suprasegmental vibrotactile speech cues has been shown to improve intelligibility. Whether this benefit stems from strengthened synchronization to the low-frequency rhythms of the target speech in the auditory cortices (i.e., cortical tracking of speech, CTS) remains unclear. Moreover, it is unknown if this multisensory enhancement is associated with strengthened coupling between auditory cortices and high-level associative areas compared to unimodal, auditory-only conditions. In the current study, magnetoencephalographic (MEG) data from 30 normal-hearing listeners was recorded while they attended to naturalistic, connected speech presented in quiet or embedded in a multi-talker background. Speech was presented either unimodally (auditory-only condition, A) or bimodally, with two types of supplemental non-auditory input: the original video of the speaker's face (audio-visual condition, AV) or speech-derived vibrations presented to the left hand either synchronously (congruent auditory-tactile condition, ATc) or asynchronously (incongruent auditory-tactile condition, ATi) with the speech signal. To further assess how non-auditory modality-specific cues shape CTS independently from the auditory input, participants were exposed to visual-only (V-only) and tactile-only (T-only) conditions. Across conditions, comprehension was assessed using yes/no questions, while CTS was quantified through "speech-brain" coherence at phrasal (0.2–1.5 Hz) and syllabic (2–8 Hz) levels. In multi-talker background noise, comprehension accuracy significantly improved in ATc conditions (vs. A), which corresponded to an enhanced syllabic CTS at the right supratemporal cortex. This improvement was associated with increased functional connectivity from this auditory region to the ipsilateral angular and inferior temporal gyri, alongside reduced connectivity with the precuneus. Contrastingly, results did not uncover significant behavioural and neural effects during AV conditions. These findings reveal that in cocktail-party conditions, speech-based vibrotactile input not only enhances comprehension and CTS of attended speech but also promotes broader synchronization with high-level, multisensory areas implicated in lexico-semantic processing. Additionally, they underscore the crucial role of syllabic CTS in supporting speech comprehension in adverse auditory environments. Altogether, these results provide novel insights into the neurophysiology of audio-tactile speech enhancement and emphasise the potential of vibrotactile stimulation linked to attended syllabic rhythm as an alternative sensory cue for speech-in-noise perception.

ID: 277

Abstract

Keywords: Statistical learning, firing rate adaptation, Auditory cortex, Inferior colliculus, cell types, Neuropixels, mouse

Role of cortical and subcortical regions in detection of sound statistics

Irene Onorato[1], Livia de Hoz[1], David McAlpine[1,2]

[1]Charité – Universitätsmedizin Berlin, Neuroscience Research Center, Berlin, Germany; [2]Macquarie University, Hearing Research Centre, Sydney, Australia; irene.onorato@charite.de

Our auditory system is incredibly good at accurately representing stimuli that rapidly change in amplitude and frequency. Neuronal adaptation has been proposed as a key mechanism for enhancing the discrimination between salient foreground sounds and background noise. This is achieved through dynamic adjustments in neuronal response sensitivity based on the statistical structure of the sound environment.

A critical aspect of this statistical learning involves the interaction between the feedforward stream and cortical feedback, forming “listening loops” within these auditory regions. To investigate this system, we performed simultaneous Neuropixels recordings from the inferior colliculus (IC) and primary auditory cortex (A1) in both anaesthetised and awake mice.

To study statistical learning, we played a continuous broadband noise stimulus for over 10 minutes. The noise intensity is modulated within a given statistical distribution that switches every 5 seconds (trial length) between two alternated contexts.

In anaesthetised animals, IC neurons exhibited short integration times, adapting quickly to the mean sound intensity within each trial but showing little influence from sound statistics across trials.

In contrast, in the awake state, neurons in both IC and A1 were less influenced by immediate trial-specific statistics and more by the global statistical structure of the protocol. This suggests a reduced sensitivity to immediate context and an enhanced capacity for a longer integration time of the sound.

Strikingly, A1 neurons in awake animals not only adjusted their average firing rates across trials but also refined their firing patterns, becoming more selective for specific noise intensities over time. This indicates an experience-dependent refinement in sensory representation that was not observed under anaesthesia.

We also found substantial inter-animal variability in the degree of adaptation and in the dynamics of firing pattern changes in A1. Our findings suggest that these differences correlate with the animals’ behavioural state, as assessed through video-based tracking of pupil size and facial movements.

To investigate the cell-type-specific contributions to statistical learning, we applied waveform classification and opto-tagging in VGAT transgenic mice to distinguish excitatory from inhibitory neurons in A1 and IC. This approach allowed us to examine the distinct roles of these cell types in encoding background sound statistics across auditory regions.

In conclusion, we explored how different neuronal populations in both cortical and subcortical auditory areas encode sound statistics and how these processes are modulated by behavioural state. These findings contribute to a deeper understanding of how the auditory system processes complex, noisy environments and may offer insights into the mechanisms underlying listening dysfunctions.

ID: 278
Abstract
Keywords: Ferret, Auditory localization, head movment, naturalistic, Behavior

Head movement and behaviour of ferrets during naturalistic auditory localization

Amit Khandhadia[1], Stefan Zigic[1], Damien Wallace[2], Kay-Michael Voit[2], Jason Kerr[2], Jennifer Bizley[1]
[1]University College London, United Kingdom; [2]Max Planck Institute for Neurobiology of Behavior; amit.khandhadia.17@ucl.ac.uk

Natural auditory behaviours require listeners to locate sounds during movement and within complex soundscapes. However, many experiments examining auditory localization behaviour often present sounds from static locations at the edge of an experimental enclosure, while requiring the subject to listen from a fixed central location. Similarly, how mammals use head movements to conduct these localization behaviours during locomotion or to a shifting sound source remain mysterious. To better understand how 3-dimensional head movements support sound perception, trained ferrets to localise sounds in a “speaker grid” environment. Within this speaker grid, forty individually controlled speakers were located within a 2m by 4m arena, placed beneath an acoustically transparent mesh floor. We trained ferrets to first come to a start position before they then had to localize a stream of 100ms noise bursts played at irregular 150-500ms intervals from one speaker location. The subject had to move to the location of the sound and remain within circle of diameter 22cm centred over the speaker for 200ms. This triggered a second speaker location to become active, and the ferret then had to localize a sound at the second speaker location, maintaining position over this second speaker for 1800ms, before obtaining a reward at a separate reward location. This closed-loop task was controlled using Deep Lab Cut live to track the position of the animal and determine its proximity to start, speaker and reward locations. Animals rapidly learned to perform this task, and one animal has been implanted with neuropixels for recordings, as well as to allow head mounted infra-red tracking. We will conduct analysis of head pose during locomotion using 8 cameras capturing high speed videos of the ferrets as they perform this task with a fixture of infrared lights attached to their heads. From these data, we can determine the roll, pitch, and yaw of the head to determine how subjects utilize their heads to locate these sounds. We hypothesise that each trajectory will be formed of a distinct sequence of movements that allow animals to initially orient, direct their movement, and when necessary, correct their course during approach to target locations. Together, these data will elucidate how animals use their head movements to actively sense and localize sounds.

ID: 279
Abstract
Keywords: developmental plasticity, critical period, electrophysiology, sensory processing, white noise

Intensity-tuned neurons in the primary auditory cortex are critical for processing complex sounds

Tommaso Zeppillo, Florian Studer, Tania Rinaldi Barkat
Department of Biomedicine, University of Basel, Switzerland;
tommaso.zeppillo@unibas.ch

Neuronal responses to sound intensity in the primary auditory cortex (A1) indicate complex encoding mechanisms that contribute to perception and behavior. Many neurons exhibit a non-monotonic firing rate response to increasing sound intensity, where higher intensities do not lead to larger neuronal responses. However, it remains unclear (1) whether A1 neurons are organized based on sound intensity preferences and how they respond to amplitude-modulated sounds (AMS); (2) whether they play a role in the perception of intensities; and (3) whether these responses are shaped by experience during developmental windows known as critical periods, during which neuronal plasticity is highly enhanced. To address these questions, we recorded A1 neuronal activity in response to sounds (4–64 kHz white noise) presented at intensities ranging from 10 to 80 dB SPL, and AMS ramping between 15-75 or 75-15 dB SPL (up or down AMS, respectively). Our data reveal distinct clusters of neurons tuned to specific sound intensities that are not tonotopically arranged, but preferentially organized within the cortical column. Cells tuned to quieter sounds exhibit distinct responses to up vs. down AMS, while those tuned to louder sounds show no difference. These distinct responses arise in A1 as they are not observed in the auditory thalamus. The role of A1 neurons for the perception of intensity was assessed by training mice to discriminate between two non-intensity-modulated sounds (NAMS) at 35 and 75 dB SPL, followed by optogenetic inactivation of A1. Behaviorally, mice efficiently discriminate NAMS; however, cortical inactivation does not affect performance, indicating that A1 neurons are not essential for non-modulated intensity perception. To explore a potential critical period for AMS, mice were exposed to 500 ms down AMS (75–35 dB SPL) during specific developmental windows, and A1 responses in adult mice were compared with non-exposed controls. We found that exposure during postnatal days P31–P38, but not during P16–P31 or P36–P50, abolishes the difference in responses between up and down AMS. In conclusion, although A1 neurons are not required to discriminate NAMS, they exhibit an intensity-tuned functional organization, whose distinct responses can be modified by exposure from P31 to P38, a critical period for AMS. Ultimately, unveiling the mechanism of intensity processing will enhance our insight into complex sound processing, crucial for language perception.

ID: 280
Abstract
Keywords: pitch, population decoding, ferret, electrophysiology

Harmonic and temporal pitch cues are encoded in distinct but overlapping populations of auditory cortical neurons

Veronica M. Tarka[1], Ben D. B. Willmore[1], Andrew J. King[1], Quentin Gaucher[2], Kerry M. M. Walker[1]
[1]Department of Physiology, Anatomy and Genetics, University of Oxford, UK;
[2]Sorbonne Université, CNRS, Inserm, Institut de la Vision, F-75012 Paris, France;
veronica.tarka@wolfson.ox.ac.uk

Pitch, the tonal quality of sound, is crucial to our experience of musical melody, enhances vocal communication, and helps us focus on one voice in a crowded room. We perceive the pitch of a periodic sound at a single fundamental frequency (F0), and this F0 can be derived from either the regular spacing of harmonic frequencies or the repetition rate of the sound's waveform in time. While previous studies suggest that auditory cortex may play a key role in processing pitch, it remains unclear how neurons in this region extract F0 from harmonic and temporal features of complex sounds. We performed Neuropixels recordings of single-neuron spiking activity in the auditory cortex of 4 anaesthetised ferrets while presenting a variety of pitch-evoking sounds that varied in harmonic composition and temporal regularity. We find that F0 is linearly decodable from neural population activity across all sound types presented, but these population codes differ depending on the acoustic content of the sound. Some neurons play a key role in encoding F0 across many different sound types, while other neurons are only informative about a sound's F0 when specific acoustic cues are present. Single neurons with consistent F0 tuning across a variety of pitch cues are often key to population-level F0 encoding across different complex sounds. In these neurons, F0 tuning does not generally correspond to their pure tone tuning. This is recapitulated at the population level, where we find that a linear decoder trained on pure tone responses cannot accurately predict the F0 of complex sounds. This suggests that subsets of auditory cortical neurons may be specialized to represent the pitch of complex harmonic and temporal pitch cues using mechanisms that cannot be predicted from their pure tone frequency preferences.

ID: 281
Abstract
Keywords: auditory scenes analysis, naturalistic scenes, semantic modelling, RSA, MEG

Learned semantic associations from real-world events co-occurrences influence the auditory perceptions of naturalistic sounds scenes.

Giorgio Marinato[1], Christian Ferreira[1,4], Michele Esposito[2], Marie Plegat[1,3], Elia Formisano[2], Bruno L. Giordano[1]
[1]Institut de Neurosciences de la Timone, CNRS, Aix-Marseille Univ., Marseille, France; [2]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands.; [3]Institut de Neurosciences des Systèmes, Inserm, Aix-Marseille Univ., Marseille, France; [4]Laboratoire d'Informatique et des Systèmes, Aix Marseille Université, UMR 7020 CNRS, Marseille, France; giorgio.marinato@univ-amu.fr

Humans naturally integrate acoustic details with higher-level contextual knowledge when interpreting the rich acoustical structures encountered in daily life. While prior research has suggested that statistical regularities influence our auditory perception through predictions, internalized schemas and attention (Southwell et al., 2017; Woods & McDermott, 2018), our focus shift instead on the semantic role (Giordano et al., 2023) of the contextual regularities. Specifically, we explore how the typical co-occurrence of sounds (e.g., birdsong with rustling leaves, traffic noise with car horns), may influence our representation and understanding of the auditory scenes.

To investigate whether such co-occurrence statistics shape neural responses to complex sound mixtures, we considered a large-scale corpus of machine-generated descriptions of real-world scenes (Auto-ACD; Sun et al., 2023) to build semantic models of sound-event relationships, including both Word2Vec embeddings and frequency-based co-occurrence statistics. We tested these models against magnetoencephalography (MEG) data, recorded while participants listened to 10 seconds long, naturalistic sound scenes, comprising common environmental and man-made sounds.

For each sliding time window of 1s in steps of 0.5s of the sensor-level MEG signal, we computed representational dissimilarity matrices (RDMs), and assessed their alignment with model-based predictions using representational similarity analysis (RSA). Results revealed modest, yet temporally structured correlations, between model-based and neural RDMs. Both Word2Vec and partial mutual information (PMI) co-occurrence models showed significant RSA correlation peaks during mid-to-late time windows, suggesting a potential influence of semantic structure on auditory scene processing. These effects were more prominent between 2 seconds and 4 seconds and 8 to 10 seconds, aligning with the hypothesis that contextual integration in auditory perception unfolds over time.

While the overall effect sizes were modest, the findings do suggest that listeners' neural representations of complex scenes are shaped not solely by acoustic features, but also by learned semantic associations. Importantly, the co-occurrence model captured aspects of neural dynamics comparable to those of Word2Vec embeddings, hinting at the brain's sensitivity to semantic regularities in real-world events.

Our study helps framing the research on top-down and bottom-up processes in auditory scenes perception, exploring how even relatively abstract semantic models can partially explain brain activity patterns of listeners. Future work should explore more refined semantic structures, task-driven modulations, and the integration of

multimodal context to deepen our understanding of real-world sound processing in the human brain.

ID: 282
Abstract
Keywords: prefrontal cortex, fmri, human, tonotopy, language

Mapping functional specializations for auditory cognition in human prefrontal cortex

Abigail Noyce[1], Wusheng Liang[1], Madhumitha Manjunath[1,3], Sahil Luthra[1,2], Barbara Shinn-Cunningham[1]
[1]Carnegie Mellon University, United States of America; [2]Stony Brook University, United States of America; [3]Boston Children's Hospital, United States of America; anoyce@andrew.cmu.edu

Prefrontal cortex (PFC) controls how we hear, especially when cognitive load is high. Whenever listeners extract complex concepts from acoustic input, remember sounds and use them to guide behavior, or merely attend to one of multiple competing sources, PFC structures are recruited. However, the organization of functional subregions within PFC is still poorly understood, in part because group-average neuroimaging methods struggle with fine-scale structure in such highly plastic brain areas. Our previous work identified five bilateral auditory-biased structures within human PFC by contrasting auditory versus visual working memory (WM) using fMRI (Noyce et al., 2022), but it is not known whether or how these areas differ in their functional specialization and contributions to auditory cognition. Here, we test whether bilateral auditory-biased structures in caudomedial superior frontal gyrus (cmSFG), transverse gyrus intersecting prefrontal sulcus (tgPCS), caudal portion inferior frontal sulcus and gyrus (cIFS/G), anterior central operculum (aCO), and frontal operculum (FO) are also (1) tonotopic, (2) language-selective, and (3) sensitive to spatial location. We first localized these structures in individual subjects (N = 17) using a direct contrast of auditory WM (animal sounds) versus visual WM (face photographs). For each structure, we tested its pitch selectivity in a phase-encoded tonotopy paradigm with an embedded 2-back WM task. We observed strong tonotopic organization in auditory cortex, but a profound absence of any pitch selectivity in any PFC structure, nor in whole-PFC maps. We then tested each region's selectivity for language processing by contrasting blocks in which written English sentences were presented word-by-word at center screen against matched blocks of non-words. We observe left-lateralized language selectivity in the more dorsal structures (cmSFG, tgPCS, cIFS/G), but no language selectivity along the operculum (aCO, FO). Whole-PFC maps of language selectivity suggest that language processing overlaps auditory-biased WM structures, but also recruits areas adjacent to those regions. Finally, we tested each region's recruitment during matched WM tasks in which the relevant feature was either the spatial locations of the stimuli or the spectral structure of the stimuli. We observe strong bilateral recruitment in both tasks, but no preference in auditory-biased PFC for one versus the other memory feature, nor do we observe contralateral biases for left-lateralized or right-lateralized stimuli. These findings lead toward a reconciliation among competing accounts of the functional organization of PFC, and a deeper understanding of its role in human hearing.

ID: 283
Abstract
Keywords: audio-visual enhancement, speech perception, neural encoding, ageing, cognition

Audio-visual enhancement of speech perception across the lifespan: evidence from behavioural and EEG data.

Mathilde de Kerangal[1], Enrico Varano[2], Jonas Auernheimer[3], Tobias Reichenbach[3]
[1]Ear Institute, University College London, London, United Kingdom; [2]Department of Bioengineering, Imperial College London, London, United Kingdom; [3]Department Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany; m.kerangal@ucl.ac.uk

Hearing is essential for communication, but as people age, understanding speech in noisy environments becomes challenging. Age-related hearing loss contributes to this phenomenon, but difficulties arise even in those with limited hearing loss. The alteration of brain mechanisms involved in the analysis of complex auditory scenes, such as attention, memory, or inhibition of irrelevant information, likely plays an important role.

Speech perception can be enhanced when the listener can both hear and see their interlocutor. Various types of visual signals can promote audio-visual (AV) enhancement, but it remains unclear whether age and impaired cognition can interfere with the neural processes that underpin AV integration.

The first aim of the current study is to investigate how age and cognitive factors affect AV integration when listeners are exposed to both natural and degraded visual speech signals. The second aim is to examine whether the neural processing of continuous speech is modulated by AV integration and ageing.

Behavioural and EEG experiments (N=33, age range: 20-72) were used for each of these purposes respectively.

Using the AV GRID corpus, two degraded versions of the videos were presented to the participants in addition to the unprocessed (natural) version. The target speaker was presented along with a two-voice babble at a signal to noise ratio corresponding to the speech reception threshold (SRT) of each participant. The subjects also completed several cognitive tasks targeting working memory (digit span and Corsi task), sustained attention (reaction time task), and inhibition (Stroop task).

As expected from previous studies, the results from the cognitive tasks correlated with age. The SRTs also correlated with age, working memory, and the Stroop scores.

The behavioural results from the AV speech task demonstrated a gradual improvement of speech comprehension with the addition of details in the visual signals.

AV enhancements did not correlate with age or any cognitive measures, indicating that the individual differences could not be attributed to these factors. This indicates that AV integration for speech remains stable with age, suggesting that it serves as a robust mechanism across the lifespan.

The neural encoding scores in the AV condition correlated with behavioural AV results, but these neural encoding scores did not correlate with age or cognitive measures.

To conclude, AV benefit was observed in the presence of natural and degraded speech signals, with greater visual details leading to significantly stronger enhancement. AV integration remained stable regardless of age or cognitive abilities,

suggesting it is a fundamental mechanism for understanding speech in noisy environments that persists across the lifespan. Neural encodings were linked to speech comprehension but showed no association with age or cognition.

ID: 284
Abstract
Keywords: auditory thalamus, MGB, optogenetics, task execution, performance

The auditory thalamus controls stimulus-driven choices

Su Jin Kim[1], Kishore Kuchibhotla[1,2,3]

[1]Department of Psychological and Brain Sciences, Johns Hopkins University, United States of America; [2]Department of Neuroscience, Johns Hopkins University, United States of America; [3]Department of Biomedical Engineering, Johns Hopkins University, United States of America; sjkim1@jh.edu

Recent work has shown that the auditory cortex (AC) is critical for initial learning of an auditory go/no-go task with simple pure tones, but becomes dispensable at expert stages. This raises two key questions: if cortical involvement fades, which auditory structures remain necessary for task execution? And, do these structures merely relay sensory information, or do they contribute to the decision process itself? The medial geniculate body (MGB) is a compelling candidate: it possesses strong links to action and decision circuits, has been implicated as a key sensory relay, and recent findings suggest it may encode higher-order choice and reward information. Here, we use temporally precise and probabilistic optogenetics to test the necessity and functional role of the MGB in expert auditory task performance. Mice were trained to lick in response to one pure tone (S+) and withhold licking to another (S−). After reaching expert performance (two consecutive days >85% accuracy), we initiated optogenetic manipulations. Full-trial silencing of the MGB (n=6, AAV1-CAMK2a-stGtACR2, bilateral) degraded performance to chance levels (d': 2.34+/-0.49 light-off vs. 0.20+/-0.55 light-on, p<0.001), with reduced hit rates, increased false alarms, slower reaction times, and diminished response vigor—demonstrating that the MGB is required for task execution even at expert levels. To dissociate sensory relay versus decision-related roles, we selectively silenced the MGB during either the stimulus period ('stimulus-only') or the choice window (closed-loop: light-on 100 ms post-tone, light-off upon first lick). Stimulus-only silencing had modest effects—performance degraded but remained above chance (d': 1.50+/-0.62 light-on), with unchanged hit rates and a mild increase in false alarms. In contrast, choice-window silencing replicated the full-trial deficits (p=0.91, two-way ANOVA), degrading performance to chance level and with animals exhibiting impaired hit rates, elevated false alarms, slower response times, and reduced response vigor. These results suggest that the MGB does not just act as a passive relay to downstream decision circuits but instead plays an active 'switchboard' role by linking learned auditory cues to appropriate actions during task execution.

ID: 285
Abstract
Keywords: Predictive coding, mismatch, deviance detection

Intracranial high-frequency activity and oscillatory cortical dynamics of auditory prediction error propagation in humans

Alejandro O Blenkmann[1], Sabine Leske[1], Vegard Volehaugen[1], Vinicius Rezende Carvalho[1], Silvia Kochen[2], Tor Endestad[1], Robert T Knight[3], Anne-Kristin Solbakk[1]
[1]University of Oslo, Norway; [2]CONICET, Argentina; [3]UC Berkeley; alejanob@uio.no

INTRODUCTION

The neural network underlying human auditory prediction error (PE) generation beyond the auditory cortex is incompletely understood. To investigate this, we recorded intracranial EEG during an auditory task from 31 adult patients with drug-resistant epilepsy.

METHODS

Patients passively listened to a stream of tones while reading. We used the Optimum-1 paradigm [1], consisting of 300 standard tones interleaved with 300 randomly presented deviant tones per block (3-10 per patient). Deviant tones differed from standards in: INT-D (intensity down), FREQ (frequency), LOC (sound source location), or TIME (shorter duration or silent gap). Electrode coordinates were derived from MRI and CT images and projected onto a standardized brain surface.

Non-epileptic channels were bipolar-referenced (N=2050), and theta (4-8Hz), alpha (8-12Hz), beta (12-30Hz), and high-frequency band activity (HFA, 64-256 Hz) power time courses were computed. Responsive channels were isolated by comparing active (0-400ms) vs. baseline (-100-0ms) periods. We used Linear Mixed-Effects models to examine the impact of the different deviant types on the oscillatory and HFA at cortical surface vertices. Statistical significance was assessed using a permutation-based approach ($p < 0.01$, FDR-corrected). Only the most significant effects are reported.

RESULTS

All deviants elicited the strongest HFA responses in a “Core PE Area” (CPEA), at the intersection between the posterior operculum, planum temporale, and temporo-parietal-occipital junction (~100-250ms, earliest ~50ms). The superior temporal plane (STP) showed increased HFA for all deviant types, most prominently for FREQ. HFA showed specific activations in the superior parietal cortex for LOC (~300ms), inferior frontal for FREQ (~200ms), and precentral gyrus for TIME.

The CPEA exhibited increased theta and alpha power (starting ~50ms). Beta power in the STP was reduced for all deviants after ~200ms. Alpha power increased early (~100ms) for INT-D in STP, and decreased after ~200ms for the other deviants (strongest for FREQ).

In the prefrontal cortex, alpha and theta power increased for all deviants (earliest ~100ms) except FREQ. Theta power decreased in STP after ~200ms for FREQ, LOC, and TIME deviants.

CONCLUSIONS

HFA responses in CPEA suggest this is the principal generator of PE, located outside the primary auditory cortex.

HFA PE propagation for “what” deviants tended to follow the auditory ventral pathway, whereas “when and where” PE tended to follow the dorsal pathway [2].

Beta suppression to FREQ deviants observed in STP is consistent with previous reports [3]. However, the overall spatiotemporal dynamics of oscillatory activity underlying the propagation of PE exhibited a complex pattern that varied across brain areas and deviant types.

[1] Näätänen et al., 2004, DOI:10.1016/j.clinph.2003.04.001

[2] Rauschecker & Afsahi, 2023, DOI: 10.1002/cne.25560

[3] El Karoui et al., 2015, DOI:10.1093/cercor/bhu143

ID: 286
Abstract
Keywords: LFP, primary auditory cortex, mutual information

Temporal Dynamics of Functional Connectivity in Awake Cat Auditory Cortex is Disrupted by Acoustic Stimulations

Xiaohan Bao[1], Carina J Sabourin[1,2], Ying Huang[3], Stephen G Lomber[1,2]
[1]Department of Physiology, McGill University, Montreal, Canada; [2]Department of Psychology, University of Texas at Arlington, Arlington, United States; [3]Research Group Comparative Neuroscience, Leibniz Institute for Neurobiology, Magdeburg, Germany; xiaohan.bao@mail.mcgill.ca

Neuronal populations in the auditory cortex encode sensory inputs by operating as an interconnected network. Sequential impulses of acoustic stimulus are encoded by neurons in cat primary auditory cortex (A1) with a temporal precision of 20~50 msec. Therefore, it can be speculated that the functional connectivity (FC) among A1 neurons can be altered at a similar time resolution, which gives rise to an intracortical network undergoing dynamic changes. As multichannel recordings at different spatial scales become increasingly popular in human and animal models, more and more quantitative approaches have been developed. In this study, we quantified FC as a function of time using mutual information (MI), which is known for its strength in better representing non-linear characteristics in neurophysiological data. We hypothesized that acoustic stimulus has a short-latency impact on the FC and the neuronal networks in A1 that can be characterized with cross-channel MI. In alert cats, we recorded local field potentials (LFPs) from 32-channel electrode matrices that were chronically implanted within the left A1. Relationship between the LFP activities from any two different recording channels was quantified by an MI value as a measure of FC. Collectively, pairwise cross-channel MI values served as a characterization of the neuronal networks in the recorded cortical area, which spanned about one-third of the surface area in A1. Neuronal activities were driven by acoustic stimulus, which consisted of multiple clicks in each recording trial. In addition to periodic and aperiodic clicks, we also presented tone pips, conspecific vocalizations, and light flashes in some recording sessions as well. At varying stimulus lags, MI values were compared, for both individual channel pairs and as in a matrix, to a baseline regardless of stimulus lags. A preliminary analysis in 4 electrode matrices showed consistently that click stimulus disrupted FC at stimulus lags smaller than 50 msec. For these specific lags, individual pairs of channels demonstrated drastic changes in the joint distribution of LFPs from the baseline, as indicated by elevated Chi-square statistics. Similarly, pairwise MI matrices were re-ordered as compared to the baseline, which was revealed by a decrease in spearman correlations at similar stimulus lags. Our current results suggested that aperiodic click stimulations disrupt baseline FC in cat A1, which may in turn impact neuronal networks and, consequently, cortical desynchronizations. Next, phase information derived from narrowband LFPs can be incorporated to examine FC at different time scale. The effect of acoustic periodicity, carrying frequency, naturalistic familiarity, and cross-modal modulations will also be investigated. This study will provide deeper insight into the temporal dynamics of cortical connectivity in A1 and pave the way for understanding human auditory cortex using electrophysiological techniques (such as MEG and ECoG).

ID: 287
Abstract
Keywords: Temporal integration, Music and speech, iEEG, Temporal Context Invariance

Human auditory cortex integrates across similar timescales in music and speech bilaterally

Zehua Kcriss Li
University of Rochester Medical Center, United States of America; zli142@ur.rochester.edu

Integrating information across the complex temporal structure that defines music and speech is one of the central challenges of human hearing. Many theories propose distinct neural integration timescales for music and speech, reflecting their differing acoustic and temporal properties. However, no studies have directly estimated integration windows for speech and music in the human auditory cortex, in part due to the challenge of measuring integration windows from nonlinear systems like the brain using coarse, non-invasive neuroimaging methods. To overcome this challenge, we measured integration windows from the human auditory cortex using spatiotemporally precise intracranial recordings from neurosurgical patients coupled with a recently developed method for estimating integration windows from nonlinear (and linear) systems (the Temporal Context Invariance (TCI) paradigm). The TCI paradigm identifies the shortest segment duration for which neural responses remain invariant across varying contexts and does not make any assumptions about stimulus features that underlie the response or the stimulus-response mapping (e.g., linear or nonlinear). Consistent with prior findings, we observed that neural integration windows substantially increase as one ascends the cortical hierarchy from primary to non-primary auditory cortex bilaterally. However, we find that integration windows were very similar for speech and music stimuli across both primary and non-primary auditory cortex in both the left and right hemisphere. These findings suggest that neural integration windows do not change substantially with the category of sound, and thus that information in music and speech is integrated using similar temporal windows.

ID: 288
Abstract
Keywords: Tracking, prediction, spatial, temporal, integration

Joint spatial and temporal processing of moving sounds

Saira Jameela Hübenette, Anne-Kristin Solbakk, Alejandro Omar Blenkmann, Anne Danielsen, Tor Endestad
RITMO Centre for Interdisciplinary Studies in Rhythm, Time, and Motion, University of Oslo (UiO), Norway; sairamh@uio.no

Tracking a moving sound requires knowledge of its location at any given time point. This implies that we must combine information about time and space to accurately predict the trajectory of the sound. Previous studies suggested that space and time for auditory stimuli are processed separately in the brain. For example, auditory spatial processing appears slower than temporal processing. Moreover, while "what" processing is strongly mediated by temporal predictability, such that the mismatch negativity (MMN) to "what" deviants is absent during temporally uncertain tasks, the MMN to "where" deviants is still present. However, despite their seemingly distinct processing pathways, they must integrate at some processing stage.

Numerous studies have investigated visual and audio-visual tracking in space, but few have studied purely auditory location tracking. The temporal and spatial aspects of auditory tracking are thus not well understood. It remains unknown if these processes are initially working in parallel and then converging at some point, or if there are several instances of convergence throughout.

In this study, young adults listened to trials consisting of a series of 5 identical tones coming from 5 different speakers placed in a semicircle in front of the person. In standard trials, the sound series started from the left or right speaker, and played with 500 ms increments towards the opposite side. Deviant trials could be either spatially, temporally, or spatio-temporally deviant. For spatially deviant trials, the 5th tone ("target") played from the first speaker of the series again, thus deviating from the spatial pattern. The temporally deviant tone played 250 ms earlier than expected, and the third type was both spatially and temporally deviant. In some trials, the last tone deviated in pitch, and participants indicated if the pitch of this tone was different from the prior ones by pressing a button with their index finger ("same") or middle finger ("different"). They changed response hand between trial blocks to control for laterality (i.e., responded with each hand for half the trials).

Behavioral results (n = 19) showed significantly slower reaction times for the temporally and spatially deviant tones, and even slower for the spatio-temporal deviants, compared to the standard trials.There was also an effect of response hand and location of expected target, in that participants responded faster when the hand and target side corresponded. EEG results (n = 9) from a subsample showed enhanced MMN and P3 amplitudes for temporal and spatial deviants compared to standard trials, and even more amplified responses for the spatio-temporal deviants, especially at frontal and central electrodes. This suggests that spatial and temporal information might be integrated at the time of responding, such that a deviation of both expected location and time has a cumulative effect.

ID: 289
Abstract
Keywords: speech prosody, cue weighting, functional connectivity

Cue weighting and frontotemporal functional connectivity during speech perception in tonal and non-tonal language speakers

Magdalena Kachlicka[1], Kyle Jasmin[2], Kazuya Saito[3], Fred Dick[3], Adam Tierney[1]
[1]Birkbeck University of London, United Kingdom; [2]Royal Holloway, University of London, United Kingdom; [3]University College London, United Kingdom; m.kachlicka@bbk.ac.uk

Listeners' weighting of perceptual cues or dimensions is not only influenced by their perceptual abilities, but also the usefulness and reliability of those dimensions for the task at hand. Previous research has suggested that such perceptual weighting is linked to functional brain connectivity between frontal and perceptual regions. For instance, when individuals diagnosed with amusia (tone deafness) perform prosody categorization, they rely less on pitch contour cues and show decreased connectivity between multiple left frontal and right superior temporal regions (Jasmin et al. 2020). Amusics' reduced weighting of pitch contour corresponds to less tightly coupled activity across these regions reflecting the fact that amusics do not integrate pitch effectively into task planning. In this study, we test the hypothesis that 'experts' in pitch contour use - namely tonal language speakers, who in their native language use pitch contours to distinguish otherwise identical words - would show higher pitch contour weighting and increased connectivity between pitch and language-related regions, when such pitch cues are present. Here we had second language learners of English, either native speakers of tonal (Mandarin and Cantonese) or non-tonal languages (e.g., Spanish, Italian, Polish) (N=20/group) undertake an fMRI task, where, they first saw a written sentence, then heard two spoken versions of the phrase. Participants then responded via button press indicating which of the two spoken phrases best matched the written sentence. Both spoken versions contained the same words, but their pitch contour and relative duration cued the placement of a phrase boundary in two different ways: in a Conflicting manner (where pitch and duration dimensions suggested a different interpretation of the stimulus) or a Consistent manner (where both dimensions suggested the same interpretation). We predict that tonal language speakers will exhibit enhanced functional connectivity between left prefrontal language and attention-related and right hemisphere pitch-sensitive brain regions, in line with behavioural evidence showing the relative up-weighting of pitch over duration in tonal language speakers (Jasmin et al., 2021; Petrova et al., 2023, Kachlicka et al., 2024). By examining how second language speakers integrate perceptual cues, this ongoing study will provide insight into the neural mechanisms of experience-driven differences in cue weighting strategies.

ID: 290

Abstract

Keywords: environmental sounds, similarity, representativeness, identification

A selection of environmental sounds for behavioural and neuroimaging research: Introducing the EnviSounds dataset

Magdalena Kachlicka[1], Jasper van den Bosch[2], Jian Kang[3], Fred Dick[3]

[1]Birkbeck University of London, United Kingdom; [2]University of Leeds, United Kingdom; [3]University College London, United Kingdom; m.kachlicka@bbk.ac.uk

With growing interest in ecologically valid stimuli and the proliferation of recognition and classification algorithms, environmental sounds are increasingly a topic of interest and study. However, there is a lack of consensus as to which sound categories should be selected, and which exemplars should represent these categories - indeed, many existing datasets include only one exemplar per sound class. However, this is not characteristic of the actual acoustic environment and can lead to brittle and unrepresentative computational solutions and experimental results. Importantly, the existing literature provides relatively little information about how human listeners perceive the recognizability, similarity and representativeness of acoustically and perceptually varying tokens of many sound classes. Here we introduce the new Environmental Sounds (EnviSounds) dataset. It provides data from over 1000 human listeners performing different tasks (recognition, similarity judgement, goodness of fit) on 53 commonly encountered environmental sound categories of human, natural and man-made sounds, each represented by 10 different sound exemplars. To ensure ecological validity and relevance, the sound categories were selected based on the sounds' overall ecological frequency of occurrence (Ballas, 1993), as well as category representativeness. For example, while many studies include a single representative sound for animals such as cats or dogs, these animals produce a range of characteristic vocalizations. Accordingly, our dataset captures this variability by including multiple characteristic sounds produced by the same source, for example meowing, purring and hissing sounds for cats and barking, howling and growling for dogs. We provide behavioural indices of within-class similarity and goodness of exemplar for all selected samples. Furthermore, our dataset also includes normative data for recognition and identification latency, accuracy, and imageability. The dataset enables researchers to select items for experiments based on pre-defined criteria across these dimensions. With the use of this dataset, environmental sounds can be considered at various levels of complexity: sounds produced by the same sources varying in their acoustics (within-class acoustic variability) and acoustically similar sounds produced by different sources (between-category confusion). The dataset can be used for specifying training sets with individual sound sources, defining the boundaries of various categories, or providing additional variables to building prediction models.

ID: 291

Abstract

Keywords: audiovisual speech, EEG, multisensory integration, linear modeling

Unique contribution of visible articulations on speech perception in noise

Aaron Nidiffer[1], Aisling O'Sullivan[2], Edmund C Lalor[1]

[1]University of Rochester, United States of America; [2]Trinity College Dublin, Ireland; aaron.nidiffer@rochester.edu

Seeing the face of a talker improves listeners' ability to understand their speech - particularly when noise obscures the speech signal. Listeners predominantly direct their gaze toward a speaker's lips which is not surprising since the lips convey general dynamic information that is correlated with the acoustic envelope and detailed articulatory shapes which convey complementary linguistic information. Neuroimaging work has also found an enhancement of lip processing regions in visual cortex when the acoustics are missing. Together, this suggests that the lips are an important feature of visual speech which the brain exploits to assist speech processing. Yet it remains unclear whether the information that confers the improved intelligibility of noisy audiovisual speech is derived from the correlated lip dynamics or the complementary lip shape. Here we present an experiment where we have modulated the amount of facial information available to listeners as they listen to audiovisual speech in noise. We did so by selectively removing visible articulatory information and preserving the two-dimensional dynamics of lip movements with a dynamic mask tied to the lips horizontal and vertical movements. We recorded electroencephalographic (EEG) signals from participants while they watched these videos. Although both types of visual speech cues are important in different situations, our results show that when speech acoustics are degraded by noise, the behavioral benefits of visual speech depend especially on the presence of complementary visual articulatory information. Indeed, when participants can see the articulatory cues, their EEG reflects that added information: First, envelope tracking when participants saw only temporal lip cues is stronger than pure auditory speech but weaker than when they also saw articulatory information. Second, electrodes over visual cortex track visual speech categories (visemes) during visual and audiovisual speech but not when the articulations are dynamically masked. Third, and importantly we observed multisensory enhancements of responses to articulatory features and only when articulators were visible. Unique responses to speech acoustics were not changed by the presence of the dynamic mask. These results, in conjunction with prior behavioral and electrophysiological results, highlight the parallel modes of visual speech benefits and the types of listening environments they provide those benefits.

ID: 293

Abstract

Keywords: Superagers, Right Ear Advantage, Puretone Audiometry, Cognitive resilience

Right-Ear Advantage as a Peripheral Signature of Cognitive Resilience in Indian Superagers

Deepashri Agrawal, Palash Kumar Malo, Thomas Gregor Issac

Centre For Brain Research, India; deepashria@cbr-iisc.ac.in

Age-related hearing loss is associated with cognitive decline, yet variability in this relationship remains underexplored, particularly in non-Western populations. We examined 200 older Indian adults, including a subset of cognitively resilient "superagers," using pure-tone audiometry and the Addenbrooke's Cognitive Examination-III (ACE-III). Compared to age-matched controls, superagers demonstrated significantly better right-ear hearing sensitivity for 3-frequency and 4-frequency PTA ($p < .05$). No such differences were observed in the left ear. Notably, ACE-III total and subdomain scores were higher in superagers across all domains ($p < .001$). A modest but significant negative correlation was observed between right-ear thresholds (4 kHz, 8 kHz, 4PTA) and ACE-III language scores ($r = –0.22$ to $–0.28$, $p < .05$), even after adjusting for age, hypertension, and diabetes. These results suggest a lateralized auditory advantage in superagers, supporting the hypothesis that better right-ear sensitivity—especially in high frequencies—may reflect peripheral correlates of cognitive resilience in aging. To our knowledge, this is the first report of such an association in a South Asian population.

ID: 294
Abstract
Keywords: FFR, SEEG, multi-sensory integration, auditory, vibro-tactie

Tactile Frequency-Following Response measured with intra-cranial recordings in humans

Anne Sophie Grenier[1], Andréanne Sharp[1], Clarisse Gotti[1], Arthur Borderie[1], Loonan Chauvette[1], Laurence Martineau[2], Paule Lessard Bonaventure[2], Emily Coffey[3], Robert Zatorre[4], Philippe Albouy[1]

[1]Université Laval, Canada; [2]CHU de Québec, Canada; [3]Concordia University, Canada; [4]McGill University, Canada; philippe.albouy@psy.ulaval.ca

It is well established that the auditory cortex can encode the acoustic features of complex sounds, as demonstrated by cortical frequency-following response (FFR). Recent studies have suggested that vibrotactile stimulation can evoke similar responses, as observed using surface EEG. However, the neural generators of these tactile FFRs, as well as their potential integration with auditory inputs, remain unknown.

To address this question, we recorded stereo-electroencephalography data in patients with drug-resistant focal epilepsy who were presented with repeated synthesized speech syllables (/da/) under three conditions: auditory only, tactile only, and combined audio-tactile, using a multichannel vibrotactile glove.

Our results show that auditory FFRs were reliably elicited in the auditory cortex and anterior hippocampus. Interestingly, tactile FFRs emerged from a more distributed network, including somatosensory areas, the inferior frontal cortex, the anterior and posterior insula, and the hippocampus. Notably, the audio-tactile condition did not result in significant modulations in FFR phase or amplitude compared to unimodal stimulation, consistent with earlier surface EEG data, and suggesting that multimodal integration likely occurs at later stages of sensory processing. These findings provide the first direct evidence of cortical tactile FFRs and open new avenues for investigating cross-modal sensory encoding in the human brain.

ID: 297
Abstract
Keywords: human speech neuroscience, magnetoencephalography, rate normalization

Vowel Duration is Encoded Relative to the Contextual Speech Rate throughout Cortical Auditory Processing

Mara Wolter[1,2,3], Andrey Zyryanov[1,2,3], Yulia Oganian[1]
[1]Centre for Integrative Neuroscience Tübingen, Germany; [2]Graduate Training Center for Neuroscience; [3]International Max Planck Research School for The Mechanisms of Mental Function and Dysfunction; m.wolter@uni-tuebingen.de

Durational information is critical for speech sound discrimination across languages, for instance in distinguishing subtle phoneme differences such as voice onset time (e.g. g-k, b-p) or vowel duration. However, perceived duration depends not only on acoustics, but also on the contextual speech rate: a speech cue of fixed duration is perceived as longer in a fast-paced context. This robust behavioral effect offers a window into the neural mechanisms underlying the perception and contextual normalization of durational speech cues - known as speech rate normalization.

Unlike visual, tactile, and non-speech auditory contexts involving durations of up to 1 sec or longer, speech perception relies on millisecond differences that evoke distinct spatial, amplitude, and latency responses in the auditory cortex. At the same time, the auditory cortex tracks speech rate by phase-locking to the rising acoustic edges in the speech envelope. Indeed, phase-locking at the speech rate frequency in right temporal and inferior frontal regions predicts the strength of rate normalization across individuals. Yet, how contextual speech rate and durational cues interact neurally remains unknown. We hypothesized that duration is initially processed in absolute terms and normalized at later processing stages.

In our magnetoencephalography (MEG) study, we tested this using vowel duration contrasts in German (e.g., satt - ˈzat' - full vs. Saat - ˈzaːt' - seed; Rum-Ruhm; Bann-Bahn), where the contextual speech rate can shift word perception. Thirty participants listened to target words with varying vowel durations, each preceded by a fast (7.5 syl/s) or slow (2.5 syl/s) sentence. After each trial, they indicated which word they perceived. We modeled the MEG responses in source space, time-locked to vowel offset, using time point by time point linear regression with predictors for precursor rate, target word, vowel duration and all interactions.

We hypothesized that if vowel duration is encoded relative to speech rate, their neural effects would overlap in time and space, with fast rate modulating neural responses similarly to long vowel duration. Conversely, effects of vowel duration in absence of rate effects would speak for absolute duration encoding.

As expected, fast-rate led to more long vowel responses compared to slow-rate context. Neurally, rate influenced evoked responses before, during, and after duration processing, across auditory, motor and inferior frontal regions. Crucially, vowel duration effects from ~130 ms post-offset were not observed distinctly from rate effects. Instead, fast rate and long vowel duration consistently shifted neural responses in the same direction, in line with the perceptual shift.

This overlap of rate and duration effects suggests that vowel duration is encoded relative to contextual rate throughout cortical processing. That is, rate normalization is an integral part of cortical processing, including early perceptual responses.

ID: 299
Abstract
Keywords: Decision-making, Rat, Sound Lateralization, Optogenetics, Drift-Diffusion Model

Investigating the Role of Auditory Cortex in Decision-Making

Mafalda Valente, Juan R. Castineiras de Saa, Raghavendra Kaushik Archak, Alfonso Renart
Champalimaud Foundation, Portugal; mafalda.valente@neuro.fchampalimaud.org

Understanding how the Auditory Cortex (ACx) contributes to perceptual decision-making has been a long-standing challenge. While it is clearly involved in processing auditory stimuli, its necessity in fundamental auditory decisions—such as sound localization based on interaural level differences (ILDs)—remains controversial. Lesion studies in rodents suggest minimal impairment in simple localization tasks after ACx damage; however, deficits are often observed in auditory decision-making tasks, especially in tasks demanding finer perceptual discriminations or relying on variables beyond stimulus features, such as temporal expectation. Therefore, ACx may play a modulatory rather than essential role for the decision-making processes.

We used ibotenic acid lesions and transient optogenetic silencing of ACx neurons to directly examine their contribution to an ILD-based discrimination task. While lesioning the cortex yielded no substantial effects, optogenetic silencing significantly impaired decision accuracy and increased anticipatory behaviors, reflected in higher rates of fixation aborts (terminating the trial prior to stimulus onset). Somewhat unexpectedly, reaction times (RTs) during cortical silencing were shorter, contrary to our initial predictions that sensory disruption would slow decisions.

To better understand these behavioral effects, we employed a computational approach using an adapted drift-diffusion model (DDM). This model incorporates two parallel processes: a proactive (anticipatory) process, and a reactive (stimulus-driven) process of evidence accumulation. While simply increasing the anticipatory (proactive) drive can mimic the increase in anticipatory responses, it did not account for the reduced accuracy nor the decreased RTs. Our model fittings revealed that reducing the firing rate of sensory neurons, and increasing noise within the evidence accumulation process, were required to accurately reproduce the full pattern of behavioral changes. Thus, ACx silencing appears to simultaneously degrade the sensory evidence and increase anticipatory behaviors.

Our work demonstrates the strength of combining targeted neural manipulations with detailed computational modelling, providing deeper insight into how cortical areas modulate perceptual decisions, and these findings support a nuanced role for ACx in auditory decision-making, its importance not necessarily in detecting auditory stimuli but potentially in optimizing perceptual decisions. By influencing both the fidelity of sensory signals and the propensity for anticipatory responses, ACx activity contributes to robust evidence accumulation and appropriate timing of decisions, suggesting an important role by balancing internally generated and stimulus-dependent information processing.

ID: 300
Abstract
Keywords: learning and memory, signal-in-noise detection, auditory cortex, SNR, rodent

Enhancing Signal-In-Noise Detection by Learning-Induced Plasticity Mechanisms in the Auditory Cortex

Nilay Atesyakar[1], Andrea Shang[1], Kasia Bieszczad[1,2,3]
[1]Rutgers University, Department of Psychology, Behavioral and Systems Neuroscience, United States of America; [2]Rutgers Center for Cognitive Science (RuCCS); [3]Department of Head and Neck Surgery & Communication Sciences, Rutgers RWJ Medical School; nilay.atesyakar@rutgers.edu

Auditory learning experiences can enhance sound detection behavior. While it is well known that associative learning induces sound-specific neurophysiological changes in the auditory cortex (ACx) (Schreiner and Polley 2014; Weinberger 2015), it is unknown whether such plasticity offers a signal processing advantage for learned signals under subsequently challenging listening conditions. Associative learning may selectively promote listening for a remembered sound signal in noisy backgrounds via learning-induced ACx plasticity to facilitate learned signal-cued behaviors. To reveal hypothesized processing advantages of learning-induced ACx changes to signal processing in noisy backgrounds, adult male rats (n = 11) learned to press a lever to receive a water reward. Shaping was followed by training in "Quiet" on a simple tone-reward associative task (8s, 5.0 kHz, 60 dB SPL pure tone). Days required to achieve 80% correct responses was used to identify rapid (< 10 days) vs. slower (> 10 days) learners. After reaching the same high levels of asymptotic performance, a behavioral memory test was conducted to assess behavioral response specificity for the trained frequency (5.0 kHz) vs. novel tones (4.2 kHz and 5.9 kHz) in "Quiet" vs. "Noise" under different signal-to-noise ratios (SNR) to determine if noisy backgrounds would weaken signal-specific behavior. In vivo ACx multiunit recordings followed all behavioral assessments in a single acute anesthetized recording session in two parts at each ACx site: (1) pure tones (0.525–47.7 kHz, 25 ms duration pure tones, 10–60 dB SPL in 10 dB steps, presented to the contralateral ear in 6 pseudorandom repetitions) were in "Quiet" to identify ACx based on short-latency, frequency-tuned, tonotopic evoked responses; (2) the same tones presented >275ms after the onset of Gaussian white noise under different SNRs (Quiet: SNR0 vs. Noise: SNR+40, SNR+20; pseudo-random presentation in 8 repetitions). Remarkably, responses to tones in noise differed in rapid (n = 6) vs. slower (n = 5) learning rats, even though testing was done when all rats performed the task (in Quiet) equally well. Rapid learning changed sound-evoked ACx activity (nrapid = 99 vs. nslower = 67 sites relative to naive: n = 7 rats, 94 sites) in two ways: (i) a signal-specific significant increase in tone-evoked activity for the behaviorally relevant sound and (ii) a general decrease in activity during steady state noise. Statistically significant differences were more pronounced for rapid vs. slower learners and emerged with higher noise conditions. Surprisingly, signal-specific behavioral responses failed in both groups during lower levels of background noise, and recovered only for rapid learners in higher levels of background noise. This work highlights the impact of learning experiences on enhancing signal detection via a non-linear auditory cortical decoding mechanism that is established early during the acquisition of novel auditory associations.

ID: 301

Abstract

Keywords: Single Sided Deafness, Cortical Development, Auditory Steady State Response, Electrophysiology

Effect of left versus right ear early childhood deafness on cortical lateralization of auditory steady state responses

Hanne Bartels[1,2], Jan Wouters[2], Astrid van Wieringen[2], Brent Sterckx[2], Rachel Lim[1], Sharon Cushing[1], Karen Gordon[1]

[1]Archie's Cochlear Implant Lab, The Hospital for Sick Children, Toronto, Canada; [2]ExpORL, Department of Neurosciences, KU Leuven, Leuven, Belgium; hanne.bartels@mail.utoronto.ca

Objective: To investigate potential effects of pre-lingual single sided deafness (SSD) on the development of hemispheric lateralization of auditory steady state responses (ASSR). Stimuli are amplitude modulation rates of 4Hz and 20Hz, which can be associated with syllable and phonemic rate, respectively, in speech.

Hypothesis: Hemispheric lateralization of syllabic rate (4Hz) will be more affected by right-sided than left-sided SSD with opposite effects for phonemic rate stimuli (20Hz).

Methods: This study included 24 (pre)adolescents with early-onset SSD. Among these, 14 children had left-sided deafness (mean age (SD) = 10.9 (3.4) years) and 10 right-sided deafness (mean age (SD) = 13.1 (3.7) years). Hearing loss was congenital in 21 children with auditory nerve aplasia/hypoplasia as the most frequent etiology (n=11). A control group of 11 normal hearing children was included (mean age (SD) = 13.9 (3.3) years). Overall, 23 children were male and 26 were right-handed. ASSR to monaural stimulation were measured using 64-channel EEG. The carrier stimulus consisted of a one-octave band white noise centered at 1 kHz, and was sinusoidally amplitude-modulated using rates of 4Hz and 20Hz. A recording time of 15min was used per condition. Modulation depth was individually determined based on fitted thresholds and comfort levels (mean (SD) dynamic range = 54.3 (6.0) dB). Time signals of the electrodes located within the parietal-temporal and occipital regions were averaged into a left and right hemispheric channel (HC) for scalp level analysis. Response significance was assessed using a one-sample Hotelling's T^2 test, and SNR and amplitude of the mean response were computed from the complex frequency spectrum. Cortical lateralization (CL) was calculated based on the HC response amplitudes, with positive values indicating right-hemispheric dominance.

Preliminary results: Linear mixed models were used to analyze comparisons between groups, ears, HCs and modulation rates, SNR (mean (SD) = 12.1 (5.0) dB) were larger in the right than left HC (p=0.04, mean (SD) difference = 1.3 (0.75) dB) and for the 20Hz than 4Hz modulation rate (p=0.03, mean (SD) difference = 1.3 (0.74) dB). Response amplitudes were larger (p<0.001) for 4Hz (mean (SD) = 0.61 (0.37) μV) than 20Hz modulation (mean (SD) = 0.61 (0.37) μV) and decreased with age (p<0.001). CL was weighted more towards the right HC for 20Hz (mean (SD) = 9.10 (22.3)) than 4Hz (mean (SD) = 1.7 (17.6)) (p=0.046) in both groups (p=0.57) and this tended to occur more for left than right ear stimulation (p=0.07). Ongoing recruitment and next steps with source analysis will refine results to gain further insights into the effect of SSD on functional asymmetry of the auditory cortices.

Significance: Ear dependent effects of SSD on hemispheric specialization could affect complex auditory processing, explaining some of the developmental deficits in language, learning and cognition in children with SSD.

ID: 303

Abstract

Keywords: autism, binaural processing, spatial hearing, EEG, spatial release from masking

Neural Correlates of Speech Perception in Autistic Children: EEG Investigation of Spatial Cue Processing

Amirhossein Chalehchaleh[1], Talat Jabeen[2], Stephen R. Dager[3], Annette M. Estes[4], Tanya St. John[4], Giovanni M. Di Liberto[1], Bonnie K. Lau[2]

[1]School of Computer Science and Statistics, University of Dublin, Trinity College, Ireland; ADAPT Centre, Trinity College Institute of Neuroscience; [2]Department of Otolaryngology – Head and Neck Surgery, University of Washington, Seattle, WA, USA; [3]Department of Radiology, University of Washington, Seattle, WA, USA; [4]Department of Speech and Hearing Sciences and University of Washington Autism Centre, Seattle, WA, USA; chalehca@tcd.ie

Autistic individuals often experience difficulties in speech perception, particularly in multi-talker environments. These challenges may stem from atypical auditory processing, including deficits in spatial auditory processing. This study investigates whether autism-related speech perception difficulties in complex auditory scenes are linked to difficulty utilizing spatial separation cues, rather than spectral discrimination deficits. Electroencephalography (EEG) was used to measure cortical responses in 38 children (19 with autism; 19 neurotypical control, NT) during a selective attention task involving continuous speech. Participants listened to speech under three conditions: (1) a single-speaker condition, (2) a co-located condition where target speech and non-target speech (masker speech) were presented from the same location, and (3) a spatially separated condition where masker speech was presented from lateralized locations. Neural tracking of speech was analysed using forward temporal response function modelling, examining the encoding of acoustic and linguistic speech features across conditions. EEG analyses revealed a significant interaction between group and listening condition (ANOVA: $F(1,36) = 10.5$, $p = 0.002$). NT children exhibited increased neural tracking of target speech in the spatially separated condition ($p = 0.016$, $d = 0.38$). In contrast, autistic children showed a reduction in neural tracking in the separated condition compared to co-located speech ($p = 0.009$, $d = 0.53$), indicating that spatial separation in multi-talker streams disrupts speech processing in autism. A similarity analysis between single- and multi-talker scenarios indicated that spatial segregation alters the neural processing especially at the word-level. These findings provide evidence of an atypical neural mechanism underlying speech perception difficulties in autism, related to impaired spatial processing of speech. The reduced ability to benefit from spatial separation, which we measured in a realistic listening task, may contribute to auditory overload and social communication challenges in autism. These insights highlight the need for targeted interventions focusing on spatial auditory processing in autism.

ID: 304
Abstract
Keywords: marmoset, vocalization, voice patch, electrophysiology, neuropixels

High density electrophysiological recordings from the medial voice-patch in marmoset auditory cortex

Rebekah E Gilliland[1], Kevin D Johnston[1], Pascal Belin[2], Ingrid Johnsrude[1], Stefan Everling[1]

[1]University of Western Ontario, Canada; [2]Institut de Neurosciences de la Timone, France; lgillil@uwo.ca

Not all sounds are created equal – vocalizations carry meaning, affect, and identity, and the brain knows it. Voice-specific patches in the auditory cortex, analogous to the well-documented face-selective patches of the visual cortex, emphasize the evolutionary importance of these communicative sounds in both humans and non-human primates. However, a thorough understanding of the neuronal and population dynamics within these voice patches is yet to be achieved.

The small common marmoset (Callithrix jacchus) has emerged as a powerful basic and preclinical nonhuman primate model, particularly well-suited for the study of vocal communication thanks to its sophisticated vocal repertoire and elaborate social dynamics. Importantly, marmosets possess a tonotopically organized auditory cortex with core, belt, and parabelt regions, and recent fMRI studies from our lab revealed distinct voice patches within this system, analogous to those found in humans and macaques.

The aim of the current study was to investigate the neuronal dynamics of voice patches in the auditory cortex at high spatial and temporal resolution. This was accomplished using ultra-high density Neuropixels probes, which enable the simultaneous recording of hundreds of well-isolated neurons, permitting both single-neuron and population level analyses. In each recording session, the probe was acutely lowered into the medial auditory voice patch of an awake marmoset, according to coordinates identified through previous fMRI studies in our lab. The marmoset then passively listened to a collection of 204 unique auditory stimuli, consisting of marmoset, macaque, and human vocalizations, as well as non-vocal sounds and pure tones. A total of 13 recording sessions were conducted across two adult male marmosets.

We recorded from over 800 well-isolated single neurons in the marmoset medial voice patch, specifically in the temporal proisocortex and the rostromedial auditory cortex, as defined by the Paxinos marmoset brain atlas. As expected, the majority of these neurons (60.6%) exhibited responsiveness to auditory stimuli. In alignment with our previous identification of this region as a voice patch, many of the neurons (33.6%) exhibited a maximal response for conspecific calls compared to other auditory stimulus types presented. Ongoing analyses will aim to further characterize the response profile and dynamics of these neurons, as well as the population dynamics that support vocalization processing. By delving into the marmoset auditory cortex, we move closer to understanding how evolution has wired the primate brain to prioritize voice above noise.

ID: 305

Abstract

Keywords: Predictive coding, 7T fMRI, mismatch effect, early schizophrenia patients, oddball paradigm

Investigating predictive coding mechanisms in the auditory cortex of healthy controls, early schizophrenia patients and clinical high-risk group using 7T fMRI

Yulia Y. Lazarova[1,6], Clement Abbatecola[1], Lucy S. Petro[1], José M. Baldaque[2], Kate Haining[1], Pradeep Dheerendra[1], Rajeev Krishnadas[4], Federico De Martino[3], David Porter[1], Tineke Grent-'t-Jong[1,5], Peter J. Uhlhaas*[1,5], Lars F. Muckli*[1]

[1]University of Glasgow, United Kingdom; [2]University College London, United Kingdom; [3]Maastricht University, Netherlands; [4]University of Cambdrige, United Kingdom; [5]Charité Universitätsmedizin Berlin, Germany; [6]University of Minnesota, United States; yulia.lazarova@glasgow.ac.uk

Under the predictive coding account, our brains rely on internal models of the world to make predictions that guide our perceptions (Clark, 2013). These models are updated when confronted with prediction errors e.g. mismatches between expectations and inputs. Alterations of predictive coding mechanisms could be at the origin of symptoms observed in psychosis (Sterzer et al., 2018). For example, there is a difference between schizophrenia patients and controls in mismatch negativity (MMN) paradigms in response to an unexpectedly different auditory tone in a sequence of otherwise uniform tones (Erickson et al., 2016). We recorded BOLD responses in the auditory cortex of control participants, early-stage schizophrenia patients (ESZ) and participants at clinical high-risk for psychosis (CHR-P), using ultra high field 7T fMRI. We directed participants' attention either towards or away from the auditory stimulation.

We acquired BOLD fMRI data in 53 participants (27 controls, 13 ESZ, and 13 CHR-P) using Siemens Magnetom Terra 7T scanner. Trials were composed of a sequence of six 150 ms tones for a total duration of 1.5s. On 80% of trials, all tones were the same (440 Hz), but in the deviant condition the last tone differed in pitch (620 Hz). To control for attention, we also presented an unrelated visual stimulus on each trial (4 'pacman' shapes which either aligned to form an illusory Kanizsa square or not). Participants performed an irrelevant task either auditory (detecting a shorter tone in the train of tones) or visual (detecting a subtle flicker in the visual stimulus). We defined four regions of interest in auditory cortex (Moerel et al., 2014): Heschl's gyrus (HG), planum polare (PP), planum temporale (PT) and superior temporal gyrus (STG).

We performed deconvolution general linear model analysis across all cortical layers, in all groups, to acquire peak BOLD response beta values. We performed Type I ANOVA with Satterthwaite's method on these values. We found a significant main effect of the stimulus - higher responses to the deviant compared to standard tones $F(1) = 158$, $p<0.001$ – indicating the violation of predictions; and significant main effect of attention $F(1) = 11$, $p<0.01$. Across all groups BOLD responses in early auditory areas (HG) were higher compared to late areas (PP, PT and STG) but the difference between the deviant and standard tone sequence was highest in the STG $F(3) = 3.55$, $p=0.01$. We observed a significant interaction effect between attention and group $F(2) = 4.3281$, $p<0.05$), as well as group and stimulus $F(2) = 3.54$, $p<0.05$, with the CHR-P standing out from the other two groups. Preliminary laminar analyses from 7 controls indicate a mismatch response across all layers. We observe a trend in the attention effect elevating responses to the deviant tone in the superficial layers of the HG when attention was focused on the auditory modality and in the deep layers when attention was directed towards the visual stimuli.

ID: 306
Abstract
Keywords: neuroanatomy, gyrification, Heschl's gyrus, cortical curvature

Structural variation in human auditory cortex gyrification revealed by data-driven curvature-based clustering

Yinuo Liu, Tyler Perrachione
Boston University, United States of America; tkp@bu.edu

The human auditory cortex is distinguished by enormous variation in the patterns of cortical folding seen across individuals. Heschl's gyrus (HG) has traditionally been described as having three distinct gyrification patterns: a single gyrus, a common stem duplication (CSD), or a complete posterior duplication (CPD) (Marie et al., 2015). Gyrification of planum temporale (PT), lying posterior to HG, is also highly variable but less well characterized (Tzourio-Mazoyer and Mazoyer, 2017). Delineation of these folding patterns is typically based on manual inspection of brain volumes or surfaces (Morosan et al., 2001; Rademacher et al., 2001), but it is clear that many patterns of individual morphology are not well described by these three categories. A better and more objective characterization of the patterns of structural variation in human auditory cortex is important, because variation in auditory cortical morphology has variously be associated with communication impairments (Leonard et al., 2001) or auditory expertise (Golestani et al., 2011). In this study, we used an unsupervised data-driven clustering approach to characterize folding patterns of the superior temporal plane (including HG and PT) based on a large sample (n=1065) of brain scans from the Human Connectome Project. Using local measures of cortical curvature obtained from FreeSurfer reconstructions, we calculated the Gaussian similarity between all pairs of subjects for gyrification within the superior temporal plane. The resulting similarity matrix was submitted to elbow-optimized k-means clustering to identify common patterns of gyrification across individuals. In contrast to the three classic categories, this data-driven clustering suggested that superior temporal morphology was best characterized by five distinct gyrification patterns. We found two distinct variations of the single HG morphotype: (i) a single HG with evident gyrification of PT (32% of hemispheres) and (ii) a single HG with minimal gyrification of PT (26%). We also found curvature patterns consistent with (iii) classic common-stem duplication of HG (16%) and (iv) classic complete posterior duplication of HG (9%). Finally, we observed a previously undescribed curvature pattern intermediate between the classic single HG and CSD, consisting of (v) a wider HG with a notch at the anterolateral end (18%). We also describe how these morphological variations are paralleled in variation in the cortical microstructure. Overall, these results expand our understanding of the range of individual variation in the morphology of auditory cortex. Future applications of these data-driven techniques may offer new routes to understanding structure-function correspondence in human auditory cortex, including how atypical patterns of gyrification characterize neurodevelopmental or neuropsychiatric variation in language and audition.

ID: 307
Abstract
Keywords: stress, speech, neural tracking

Neural tracking of stress and meter in continuous speech

Camila Zugarramurdi[1], Eleonora Beier[2], Katsuaki Kojima[3], Srishti Nayak[4], Reyna L. Gordon[4], Cyrille Magne[5], Yulia Oganian[6]
[1]Facultad de Psicología, Universidad de la Republica, Uruguay; [2]Center for Mind and Brain, University of California, Davis; [3]Perinatal Institute, Cincinnati Children's Hospital Medical Center, Department of Pediatrics, University of Cincinnati College of Medicine; [4]Dept. of Otolaryngology - Head & Neck Surgery, Vanderbilt University Medical Center; [5]Psychology Department, Middle Tennessee State University; [6]Center for Integrative Neuroscience, University of Tübingen;
czugarramurdi@psico.edu.uy

Syllabic stress provides critical acoustic and pragmatic cues during speech comprehension. Moreover, the alternation of stressed and unstressed syllables defines speech meter, which is known to facilitate speech recognition. Acoustically, in English, stress is defined by patterns of intensity, pitch and duration. Although it has been hypothesized that stressed syllables may drive neural entrainment to the speech signal, empirical evidence for the effects of stress and meter on neutral responses to speech are scarce. Recent work has revealed that the human speech cortex tracks syllables by responding to moments of rapid increase in speech intensity – peakRate events – which occur close to vowel onsets. Based on this acoustically defined marker of syllable processing, here we examine how stress and metric regularity influence 1) evoked neural responses to syllables and 2) oscillatory phase-locking to peakRate events. We recorded EEG while participants (n=26) listened to short childrens' stories. Participants heard one story that followed an irregular,speech-like meter typical of every-day speech, and another story that followed a regular meter with alternating stressed and unstressed syllables (x - -). Using temporal response function (TRF) modeling, we isolated the effects of syllabic stress and metric regularity on neural responses to peakRate events. Both stress and meter modulated early (within 50-150 ms) cortical responses to peakRate events. Specifically, stress enhanced neural responses to stressed syllables as compared to unstressed syllables, suggesting that stress is identified early during auditory processing. Moreover, while early responses to louder syllables were enhanced overall, this effect was attenuated in the metrically regular condition, suggesting a reduced reliance on acoustic cues in favor of processing based on top-down metric expectations. Furthermore, we found that neural responses to unstressed syllables of larger intensity were temporally more extended (300-400ms) than for stressed syllables of similar intensity. In the time-frequency domain, neural phase-locking around 2 Hz (corresponding to stress rate in stimulus) was stronger for stressed than for unstressed syllables. Such phase-locking patterns have been previously interpreted as endogenous oscillatory entrainment. However, an explicit model of evoked neural responses showed that this phase-locking pattern reflects the frequency and shape of evoked responses. This result directly shows that delta-phase locking to stressed syllables does not necessarily reflect oscillatory entrainment. It also provides a novel approach bridging between analyses of evoked responses and phase-locking to continuous speech.Overall, our findings reveal that neural responses to individual syllables integrate acoustic cues with perceived stress patterns and metric expectation. These neural dynamics may underlie the perceptual benefits of metrically regular speech, such as poetry.

ID: 309

Abstract

Keywords: speech adaptation, neural tracking, time-compressed speech, EEG

Impact of time-compression and background noise on content memory recall and neural tracking of speech

Mako Ikeda, Walter Dych, Sung-Joo Lim

Binghamton University, United States of America; sungjoo@binghamton.edu

With the growing use of online education platforms, students often prefer to learn contents using accelerated playback speeds. While listeners can rapidly adapt to acoustically distorted speech signals like time-compressed speech or speech in noise, memory recall of the speech contents declines with higher time-compression rate, especially in adverse listening conditions. However, the neural bases of the memory detriment from acoustic degradation through accelerated speech and/or background noise remains unclear–that is, whether the memory detriment is rooted in the brain's inability to track speech signals. Here, we examined how speech distortions from time-compression influence the fidelity of neural tacking of acoustic features of speech (i.e., speech envelope onsets), and whether the presence of background noise exacerbates this effect. We measured the EEG of 28 participants who listened to six recorded lectures at varying time-compression rates (1.0x, 1.5x, 2.0x) and background noise levels (quiet vs. babble noise at +10 dB SNR) in a fully counterbalanced within-subjects design. We operationalized participants' ability to learn and retain the lecture content based on lecture-specific content-knowledge quizzes administered immediately following each lecture. Behaviorally, listeners' long-term memory recall of the lecture content was significantly worse with higher time-compression and in the presence of background noise. The fidelity of the neural representation of the speech signal, indexed by the speech envelope reconstruction accuracy, was significantly lower when background noise was present irrespective of the time-compression rate. Nonetheless, the impact of speech rate on the extent of neural tracking of speech, measured through the multivariate temporal response functions (mTRFs), exhibited an opposite trend between listening in quiet vs. with background noise, such that faster speech rate was associated with enhanced mTRF in quiet but diminished mTRF in noise. These results suggest both variation in and complex interactions between different sources of additional cognitive demands in speech listening, their effect on neural tracking of the acoustic features of speech, and their implications for higher-level speech comprehension. Researchers examining the effects of adverse listening factors on cortical coding of speech sounds may want to consider whether all sources of adversity affect neural tracking of speech in the same way, and what this suggests about the potential for multiple domains of cognitive operations that underlie listening effort. These results also have real-world implications in considering the use of speech acceleration for studying by students, including whether this manipulation of recorded lectures is differentially efficacious in quiet vs. noisy listening environments.

ID: 310
Abstract
Keywords: auditory cortex, vocalizations, population representations

Categorical encoding of gerbil vocalizations by auditory cortex populations

Estelle E in 't Zandt, Dan H Sanes
New York University, United States of America; eei3718@nyu.edu

Background:

Variability in the acoustic features of specific vocalizations can convey information about an individual's identity, called vocal dialects. A fundamental question is how the central nervous system represents vocal categories, while also distinguishing between the subtle acoustic differences that communicate social identity. Here, we address the issue of call categorization by analyzing the response of adult gerbil auditory cortex (AC) neurons to a large array of variants recorded from the animal's own family and those of two other families. Gerbils are a favorable rodent species to address this question because they are highly social, living in multi-generational families, and produce a rich vocal repertoire (Peterson et al., 2024).

Methods:

We first investigated the ability of AC populations to categorize 4 vocalization types in awake, freely-moving adult Mongolian gerbils (Meriones unguiculatus) (n=5; 3M, 2F). We used chronically-implanted high-density silicon probes to record wirelessly from single AC neurons while presenting a large set of variants (n=300) for each of 4 vocalization categories (n=1200 stimuli, 5 trials each). The vocalizations were obtained from overnight audio recordings of individual gerbil families, one of which was always the implanted animal's own family. We used a Euclidean distance classifier to quantify how well both call type and family identity could be decoded from AC population temporal responses.

Results:

Vocal categories could be decoded from AC populations (n=857 single units) significantly above chance (90.0% +/- 0.9; chance = 25%), despite natural levels of variance in the stimulus set. When controlling for acoustic similarity between vocalizations, we found that AC populations exhibited a boundary effect, such that pairwise neural discrimination of within-category pairs was ~10% worse than across-category pairs for 3 of the 4 vocal categories.

AC populations were also able to decode the family identity of each of the 4 call types significantly above chance (70-73% across call type; chance = 50%). However, AC populations did not overrepresent own family vocalizations, as quantified by percent of the population responsive to each family's identity.

Conclusions:

Our results suggest that both vocal category and family identity are represented by the same AC network, even for single syllables. Our results are consistent with subtle acoustic differences being sufficient to distinguish family vocalizations from one another, although the effect of prior experience has not been tested.

ID: 311
Abstract
Keywords: auditory cortex, categorization, neuropixels, electrophysiology, neurometrics

Neuronal correlates for auditory categorization and bias in mice

Anjali K. Sinha, Gozde Erdil, Jared S. Collina, Yale E. Cohen, Maria N. Geffen
University of Pennsylvania, United States of America;
mgeffen@pennmedicine.upenn.edu

The categorization of sensory stimuli is a fundamental cognitive computation and is critical for many aspects of learning, perception, and prediction. Although auditory categorization has been linked to the neuronal activity in the auditory cortex, we do not fully know how categorization is implemented at the level of neuronal circuits and how it emerges in population neuronal activity with learning.

We recently showed that during categorization learning, mice use multiple complementary strategies (Collina et al., 2025). These strategies affect the animals' perception of stimuli category under uncertain conditions, e.g., for stimuli that surround the category boundary. Additionally, animals modify these strategies based on their training. To understand neuronal correlates of auditory categorization, we recorded neuronal activity from the auditory cortex of mice using implanted high-density electrodes while they perform an auditory categorization task.

Mice were implanted with custom-made chassis housing Neuropixels electrodes in the auditory cortex and trained to discriminate between a set of logarithmically spaced 6 tones each in high- and low- frequency category in a 2-alternative forced choice (2AFC) task. While recording neuronal activity, we tested how well they categorized tones around the category boundary, and whether this was influenced by biasing the distribution of the presented stimuli toward high- or low- category tones. Preliminary analysis indicates that approximately 30% of the recorded neurons exhibited category preference, as measured by the difference in the firing rate of neurons for stimuli between categories compared to within category stimuli. Additionally, population decoding accuracy was higher for stimuli between than within categories. To examine how neuronal population reflect behavior, we compared the neurometric and psychometric thresholds across sessions. The thresholds were well aligned between neurometric and psychometric estimates, suggesting that neuronal activity reflected behavior. Moreover, a subpopulation of neurons was choice-selective, as their firing rate reflected the choice of the mouse across trials. Our preliminary results show that the neuronal activity in the auditory cortex during auditory categorization reflects behavioral performance. We are currently analyzing how the neuronal population activity evolves throughout auditory learning and how biasing the category distribution across different sessions affects neuronal activity.

ID: 312

Abstract

Keywords: Neural tracking, language comprehension, speech processing

Linguistic structure and language familiarity sharpen phoneme encoding in the brain

Filiz Tezcan[1,3], Sanne ten Oever[1,2,3], Fan Bai[1], Noémie te Rietmolen[1], Andrea E. Martin[1,3]

[1]Max Planck Institute for Psycholinguistics; [2]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University; [3]Donders Centre for Cognitive Neuroimaging, Radboud University,; tezcanfiliz@gmail.com

How does the brain turn a physical signal like speech into meaning? It draws on two key sources: linguistic structure (e.g., phonemes, syntax) and statistical regularities from experience. Yet how these jointly shape neural representations of language remains unclear. We used MEG to track phonemic and acoustic encoding during spoken language comprehension in native Dutch, Mandarin-Chinese, and Turkish speakers. Phoneme-level encoding was stronger during sentence comprehension than in word lists, and more robust within words than random syllables. Surprisingly, similar encoding emerged even in an uncomprehended language—but only with prior exposure. In contrast, acoustic edges were briefly suppressed early in comprehension. This suggests that the brain's alignment to speech (in phase and power) is robustly tuned by structure and by learned statistical patterns. Our findings show how structured knowledge and experience-based learning interact to shape neural responses to language, offering insight into how the brain processes complex, meaningful signals.

Late Breaking Submissions

ID: 324

Abstracts (late breaking)

Keywords: voice identity, mental voice representations, computational modelling, similarity rating

Core dimensions of the perceptually relevant voice-identity space

Claudia Roswandowitz[1,2], Florian Mahner[3], Martin Hebart[3,4], Volker Dellwo[2]

[1]Technische Universität Dresden, Dresden, Germany; [2]University of Zurich, Zurich, Switzerland; [3]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; [4]Justus Liebig University Giessen, Giessen, Germany; claudia.roswandowitz@tu-dresden.de

Empirical investigations on the mental representation of vocal identities are largely challenged by two factors: i) the large number of voices with which each individual is familiar, and ii) the multidimensional nature of the physical voice signal. Therefore, a realistic characterisation of voice representation should encompass a wide variety of voices and not be limited to a small set of acoustic features.

This project provides new evidence on the core dimensions of the perceptually relevant voice identity space. The study integrates data from a voice corpus containing 337 speakers, large-scale crowdsourced perceptual similarity ratings, acoustic voice analysis, and descriptive voice labelling, in order to uncover the key dimensions that shape voice identity perception.

Using computational modelling, I investigated the underlying architecture of this perceptual space, and identified interpretable dimensions that drive voice similarity judgments. The model trained on similarity ratings from 359 participants revealed a representational embedding with five core dimensions, predicting voice similarity with 76% accuracy.

I used representational similarity analysis (RSA) to interpret the five core dimensions observed by the model. I correlated a set of 15 acoustic RDMs with dimension-specific similarity RDMs. This analysis revealed that each dimension is explained by a unique combination of voice features, including well-known identity-related features such as fundamental frequency (F0), spectral source characteristics, and harmonicity measures. Notably, time-varying features, including prosody and spectral shape measures, also accounted for a significant proportion of the similarity ratings specific to each dimension. Interestingly, the dimension-specific acoustic features align with the descriptive voice quality labels provided by professional phoneticians. These findings provide novel insights into voice identity representations by focusing specifically on the perceptual relevance and interpretability of the space dimensions.

ID: 325
Abstracts (late breaking)
Keywords: speech in noise, olsa, auditory cortex, audio-visual integration, deepfake

Integration of auditory and visual stimuli in the auditory cortex

Jasmin Riegel, Alina Schüller, Tobias Reichenbach
Friedrich-Alexander-Universität Erlangen Nürnberg, Germany; jasmin.riegel@fau.de

Understanding speech in noisy environments can be challenging. Seeing the speaker provides additional cues that aid comprehension. However, it remains unclear which exact aspects of the visual signal support speech comprehension, and how auditory and visual information integrate in the auditory cortex. Here we investigate different visual signals, from simplified outline faces to AI-generated avatars and natural videos. We determine to which degree they support speech-in-noise comprehension and how they influence the neural tracking of speech rhythms in the auditory cortex.

Therefore we recorded 2 hours of MEG data from 18 normal-hearing participants while they watched videos of a speaker articulating grammatically correct but semantically unpredictable German sentences. The sentences were five-word matrix sentences from the Oldenburger Satztest. After each sentence, participants repeated the words they understood. The sentences were presented in seven conditions. There were three types of audio-visual (AV) conditions, the original video, an outline of the speakers face and an avatar. The avatar was generated by animating a still image of the participant with the respective audio. The fourth condition was an audio-only condition. A still image of the speaker was provided, while the audio was playing. The last three conditions were visual-only (V) conditions. Here the three video types from the AV condition were presented with a muted audio. In this condition the participants were advised to repeat only the first word in each sentence.

We first analyzed the behavioural data and found that all visual signals aided speech-in-noise comprehension significantly compared to the audio-only condition. The original video had the strongest effect on comprehension. Both edited video styles supported comprehension in a less intense way. Secondly, we analyzed the neural MEG data using temporal response functions (TRFs) for source-reconstructed auditory cortex activity, focusing on delta, theta, and broadband neural responses linked to speech processing. Significant neural activity was observed across all audio-only and audio-visual conditions. To assess audio-visual integration, we subtracted the audio-only response from the audio-visual conditions, revealing differences, particularly in later peaks. Additionally we compared how well the neural data was reconstructed by the models of the different condtions and in a third step we examined the relation between the neural measures and the behavioural results.

These findings suggest that AI-generated avatars can enhance speech processing by modulating neural entrainment, offering insights into audio-visual integration and potential applications for improving communication in challenging listening environments.

ID: 326

Abstracts (late breaking)

Keywords: auditory, bat, default mode network, fMRI, asymmetry

Chiropteran Neuroimaging: Structural and Functional Magnetic Resonance Imaging of the Pale Spear-Nosed Bat

Stuart Washington[1], Kyle Shattuck[2], Jan Steckel[3], Herbert Peremans[3], Elisabeth Jonckers[3], Rukun Hinz[3], Tom Venneman[3], Monica Van den Berg[3], Lisbeth Van Ruijssevelt[3], Thomas Verellen[3], Dominique Pritchett[1], Jan Scholliers[3], Sayuan Liang[3], Sönke Von den Berg[4], Stephen Savoia[5], Partha Mitra[5], Stephen Lin[1], Paul Wang[1], Marleen Verhoye[3], Annemie Van der Linden[3], Karl-Heinz Esser[4], Georgios Keliris[6]

[1]Howard University, United States of America; [2]Georgetown University Medical Center, United States of America; [3]University of Antwerp, Belgium; [4]University of Veterinary Medicine in Hannover, Germany; [5]Cold Spring Harbor Laboratory, United States of America; [6]Harvard Medical School, United States of America; sdw4@georgetown.edu

Echolocating bats live huddled together in colonies comprising hundreds of individuals and use complex sounds to communicate and to navigate. These highly social and vocal species make ideal subjects for functional magnetic resonance imaging (fMRI) studies of auditory social communication given their relatively hypertrophic limbic and auditory neural structures and their reduced ability to hear MRI gradient noise. Establishing the existence of neural networks related to social cognition (e.g., default mode-like networks or DMLNs) in order Chiroptera could pave the way towards a new frontier in the study of mammalian socialization and communication. We measured blood oxygenation level dependent (BOLD) signal at 7T from nine lightly anesthetized pale spear-nosed bats (Phyllostomus discolor). Specifically, we performed independent components analysis (ICA) and revealed 15 resting-state networks. We also measured neural activity elicited by noise ripples (on: 10 ms; off: 10 ms) that span the ultrasonic hearing range (20-130 kHz) of this species. Resting-state networks intersected parietal, occipital, and auditory cortices, along with auditory brainstem, basal ganglia, cerebellum, and hippocampus. We determined that two out of a possible four midline networks were the best candidates for DMLN. We also found two predominantly left and two predominantly right auditory/parietal cortical networks. Regions within all four auditory/parietal cortical networks have been demonstrated to respond to social calls. As expected by the emergence of side-band inhibition in the inferior colliculus, ultrasonic noise ripples significantly activated the auditory brainstem (cluster-level: $p=5.27 \times 10^{-5}$, FWE correction, kE=7613) yet deactivated the auditory/parietal cortex ($p=2.08 \times 10^{-9}$, FWE correction, kE=17452). Iterative ("jack knife") analyses revealed consistent, significant functional connections between left, but not right, auditory/parietal cortical networks and DMLN nodes, especially the anterior-most cingulate cortex. Thus, a resting-state network implicated in social cognition displays more distributed functional connectivity across left, relative to right, hemispheric cortical substrates of audition and communication in this echolocating bat species. The application of advanced histological methods to 12 ex-vivo Phyllostomus discolor brain samples that have also undergone structural imaging (i.e., T2-weighted 3D rapid spin echo and spin-echo diffusion weighted) increase the likelihood of generating detailed, 3D population-based atlases as a computerized anatomical reference for these and future chiropteran functional neuroimaging results.

ID: 327
Abstracts (late breaking)
Keywords: Mcgurk effect, musical training, perceptual history, audiovisual processing

Musical Training and Prior Experience in Multisensory Speech Perception: Neural Insights from the McGurk Effect

Jihyun Lee[1,3], Ji-Hye Han[1,3], Sungmin Jo[1,3], Hyo-Jeong Lee[1,2,3]
[1]Hearing, Balance and Integrated Neuroscience Lab, Hallym University College of Medicine, Korea, Republic of (South Korea); [2]Department of Otolaryngology, Hallym University College of Medicine, Anyang, South Korea; [3]Ear and Interaction Center, Doheun Institute for Digital Innovation in Medicine (D.I.D.I.M.), Hallym University Sacred Heart Hospital, Anyang, South Korea; jihyunlee@hallym.ac.kr

Introduction: Speech perception hinges on the integration of auditory and visual cues, shaped by distinct perceptual and cognitive processes. Among perceptual factors, musical training significantly enhances multisensory processing. In contrast, cognitive factors, such as expectations shaped by prior perceptual experiences, influence neural responses to incoming sensory information. To explore these mechanisms, we investigated cortical activity during the McGurk illusion, a classic example of audiovisual speech integration, with a focus on how musical training and perceptual history affect multisensory speech perception.

Methods: Musicians and age-matched non-musicians participated in an electroencephalography (EEG) experiment employing a McGurk task. Participants viewed congruent audiovisual syllables (/pa/ video with /pa/ audio), incongruent syllables (/ka/ video with /pa/ audio), and auditory-only stimuli (/pa/ audio alone). Five conditions were analyzed based on stimulus type and perceptual responses. The rate of illusory percepts and cortical alpha power were compared between groups using dynamic imaging of coherent sources (DICS).

Results: Behaviorally, no significant group differences were found in susceptibility to the McGurk illusion. However, musicians showed increased frontal alpha power during audiovisual integration tasks, including congruent trials and auditory responses in incongruent trials. In contrast, non-musicians exhibited increased parietal alpha power under similar conditions. Additionally, participants occasionally reported illusory percepts in response to auditory-only stimuli. Further analyses—examining responses to auditory-only stimuli following congruent or incongruent audiovisual trials—revealed no significant alpha modulation clusters in musicians, while non-musicians showed significant clusters in temporal regions.

Conclusions: These findings indicate that musical training shapes neural mechanisms underlying audiovisual speech integration. Musicians appear to utilize enhanced frontal-mediated top-down control, whereas non-musicians predominantly engage parietal and temporal regions indicative of sensory-driven processing. Moreover, musicians' reduced neural modulation during subsequent auditory-only trials suggests decreased dependence on recent perceptual history compared to non-musicians.

ID: 328
Abstracts (late breaking)
Keywords: auditory semantics; AI-based modeling, real-world audition; RSA; fMRI

Acoustic-to-Semantic Transformation of Natural Sounds: The role of Frontal Areas

Maria Araújo Vitória[1], Marie Plegat[2,3], Giorgio Marinato[2], Christian Ferreyra[3,4], Michele Esposito[1], Christian Herff[5], Bruno Giordano[2], Elia Formisano[1]

[1]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University Maastricht, The Netherlands; [2]Institut des Neurosciences de La Timone, UMR 7289, CNRS and Universite´ Aix-Marseille, Marseille, France; [3]Institut de Neurosciences des Systèmes, UMR 1106, INSERM and Universite´ Aix-Marseille, Marseille, France; [4]Laboratoire d'Informatique et des Systèmes, UMR 7020, CNRS and Universite´ Aix-Marseille, Marseille, France; [5]Department of Neurosurgery, School for Mental Health and Neuroscience, Maastricht University Maastricht, The Netherlands; maria.araujo.vitoria@maastrichtuniversity.nl

The human brain seamlessly processes and recognizes sounds from the environment, yet. the neural computations underlying the transformation of low-level acoustic information into higher-level semantic representations remain poorly understood. While prior research has detailed how attributes such as frequency, pitch and spectro-temporal modulations are encoded in the auditory cortex, the full trajectory of sound recognition - especially beyond auditory areas - has not been fully explored.

Giordano et al. showed that sound-to-event deep neural networks (DNNs) outperform both purely acoustic and semantic models in predicting responses in the superior temporal gyrus (STG), suggesting that the STG encodes intermediate "hyperacoustic" representations. However, their analysis was limited to superior temporal regions, leaving the role of other brain areas, particularly frontal regions, unclear.

In this study, we collected 7T fMRI data from 10 participants as they listened to a diverse set of natural sounds (N=300/pp sampled from N=450 total stimuli). Using representation similarity analysis, we compared the brain activity patterns in superior temporal and frontal regions of interest with representations from a sound classification DNNs (Yamnet) and a language semantic model ("word2vec") . Additionally, we introduced a novel ontological model of auditory semantics, organizing sounds based on physical production mechanisms ("how") and source identity ("who/what"). Our results reveal that correlations between brain activity with semantic/ontology models increase in posterior STG and, more prominently, in frontal areas. Notably, the ontology model based on physical properties outperformed semantic model based on sound action linguistic descriptors.

These findings suggest that frontal regions contribute to higher-order auditory computations supporting the transformation of acoustic inputs into abstract, conceptual representations. Ongoing data collection and analysis will further elucidate the computational role of these regions in sound recognition.

ID: 329
Abstracts (late breaking)
Keywords: Deafness, neuroanatomy, cross modal plasticity, auditory cortex

Brain-wide mapping of neuroanatomical connections to the auditory cortex of hearing and deaf mice

Tom Harmon, Andrew Jin, Evelyn Hardin, Richard Mooney
Duke University, United States of America; thomas.harmon@duke.edu

Cochlear implants and AAV-based gene therapy have made it possible to restore hearing in congenitally deaf humans at different stages of life. However, the total absence of auditory experience prior to these interventions is likely to alter the connectivity of the auditory system in a manner that may impede hearing restoration. In fact, prior studies using conventional tracing methods in surgically-deafened rodents and congenitally deaf cats have revealed expanded connections to the auditory cortex from adjacent non-auditory cortical regions. However, whether these approaches can provide a complete brain-wide account of altered neuroanatomical connections to the auditory cortex is uncertain. Therefore, we used an intersectional viral and genetic approach to map the brainwide neuroanatomical connections to the primary auditory cortex of adult mice that were either congenitally deaf or equipped with normal hearing. Specifically, we injected retrograde viral vectors into the primary auditory cortex of adult Ai14 reporter mice that were either homozygous for a nullifying mutation to the TMC1 locus, rendering them congenitally deaf, or hemizygous for the mutation and had normal hearing throughout postnatal development. We used generalized multivariate linear regression to estimate the influence of hearing on presynaptic neuron counts, while accounting for other factors (e.g. transfection efficiency, age, sex) that could influence neuron labeling. In deaf mice, this analysis revealed expanded neuroanatomical connections from subregions of the orbitofrontal, ectorhinal, perirhinal, insular, somatosensory, and visual cortex and reduced connections from the basomedial amygdala and most thalamic nuclei. However, neuroanatomical connections from most other regions to the auditory cortex were indistinguishable between hearing and deaf mice. In summary, we found evidence of previously unknown changes to neuroanatomical connections to the auditory cortex of congenitally deaf mice, raising new considerations for efforts to restore full auditory function to congenitally deaf patients.

ID: 330
Abstracts (late breaking)
Keywords: Mouse, Auditory, Cortex, jGCaMP8m, Widefield

Primary and secondary functional divisions of the mouse auditory cortex

Janek Peters[1], Samuel Arends[1], Andreas Knoben[2], Bernhard Englitz[1]
[1]*Donders Center for Neuroscience, Radboud University, Nijmegen, NL;*
[2]*Swammerdam Institute for Life Sciences, University of Amsterdam, NL;*
jnkptrs@gmail.com

The large-scale functional organization of the cortex is an essential component for understanding neural computation. In the mouse auditory cortex there has been substantial debate about the large-scale organization, in particular reliably determining different subfields. Previous studies therefore have often pooled data across fields or used coarse approximations, limiting the achievable insight.

We here image the entire auditory cortex using a fast and bright Ca-indicator (jGCaMP8m) at high speed (100fps), providing clear divisions of the auditory cortex on the basis of rapid response modulations, which are well-conserved between individuals and allow for data-driven alignment of cortical maps. Alongside characteristic primary response areas, we highlight that the posteroventral temporal association area (pTeA/vTeA) exhibits a tonic increase of activity specifically to longer sounds, coincident with a commensurate dampening of A2 activity. Further, on the basis of trial-by-trial analysis, we find dynamic changes in response fidelity on multiple time-scales, which differ across subfields.

The present results provide the basis for more reliable, precise and meaningful analysis of neural responses from the auditory cortex, in particular when combined with cellular imaging on the basis of the estimated fields. Additionally, single trial analysis on large-scale recordings provide a new opportunity for directly relating behavior to neural activity.

ID: 331
Abstracts (late breaking)
Keywords: auditory cortex, astrocytes, learning

The dynamics of auditory cortical astrocytes during sensorimotor learning

Sharlen Moore, Travis You, Ziyi Zhu, Kishore Kuchibhotla
Johns Hopkins University, Krieger School for Arts and Sciences, Department of Psychological and Brain Sciences.; smoorec2@jhu.edu

Astrocytes, the predominant glial cell type in the brain, have emerged as active participants in neural information processing and plasticity. However, their functional dynamics across learning remain poorly understood. Here, we used chronic fiber photometry and two-photon calcium imaging to longitudinally track populations or individual astrocyte Ca2+ dynamics in the auditory cortex of awake mice (expressing Aldh1l1-dependent GCaMP6s in an inducible manner) across the acquisition of an auditory discrimination task. We trained mice to lick to a tone for water reward (S+) and withhold from licking to another tone (S−) to avoid a timeout. We added catch trials in which the reward contingency was changed in a block-based manner to understand how these signals change upon contextual cues. Over several days of training, astrocytes exhibited learning-related modulation of their Ca2+ dynamics, with cells showing enhancement of their evoked responses during rewarded trials. This increased activity arose within the first 100 trials of reward-based training and was not due to licking as the same astrocytes exhibited suppressed activity on errors of action (incorrect licking to the S-, false alarm). We observed two timescales of astrocyte activation, one ‘early-in-trial’ that aligned with the tone and initial lick. Another appeared later-in-trial, was broader and significantly larger, potentially associated with motor, reward and post-reward processing. Omitting reward on correct trials (hits, S+) led to biphasic responses where a transient increase in activity was followed by a profound suppression, suggesting that reward consumption may drive extended increases in astrocyte activity. Receiving extra reward amount (2x), led to a potentiation of these responses. Similar patterns of activity arose in correct trials with no licking (correct rejects to the S-) in the same reward-shift block. Finally, baseline activity in astrocytes efficiently tracked the reward-context. Together, these data suggest that astrocytes are tracking a more abstract context variable irrespective of movement. Interestingly, baseline reward-related signals extended beyond individual trials, potentially playing a role in maintaining a signature of reward and trial history for upcoming choices. Our data suggest that coordinated astrocyte ensembles may provide a scaffold for integrating reward signals with sensory processing to facilitate learning, potentially bridging trial-level and inter-trial computations. This study expands our understanding of astrocyte contributions to neural circuit dynamics underlying adaptive behavior.

ID: 332

Abstracts (late breaking)

Keywords: Discrimination learning, Auditory Cortex, Cholinergic Basal Forebrain, Longitudinal two-photon calcium imaging

Role of the cholinergic system in early sensorimotor acquisition

Jennifer Lawlor[1], Sarah E. Elnozahy[4], Fangchen Zhu[1], Ziyi Zhu[2], Aaron Wang[5], Fengton Du[2,6], Tara Raam[7], Kishore V. Kuchibhotla[1,2,3]

[1]Department of Psychological and Brain Sciences, Krieger School of Arts and Sciences, Johns Hopkins University, Baltimore, MD, USA; [2]The Solomon H. Snyder Department of Neuroscience, School of Medicine, Johns Hopkins University, Baltimore, MD, USA; [3]The Johns Hopkins Kavli Neuroscience Discovery Institute, Baltimore, MD, USA; [4]Sainsbury Wellcome Centre for Neural Circuits and Behaviour, University College London, London, UK; [5]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [6]Janelia Research Campus, Howard Hughes Medical Institute, VA, USA.; [7]Department of Biological Chemistry and Department of Neurobiology, David Geffen School of Medicine, University of California, Los Angeles, CA, USA; jlawlor3@jhu.edu

During sensorimotor learning, animals link a sensory cue with actions that are separated in time using circuits distributed throughout the brain. Learning thus requires neural mechanisms that can operate across a wide spatiotemporal scale and promote learning-related plasticity. Neuromodulatory systems, their broad projections and diverse timescales of activity, meet these criteria and have the potential to link various sensory and motor components. Yet, it remains unknown the extent to which this proposed model of plasticity occurs in real-time during behavioral learning. The acquisition of sensorimotor learning in a go/no-go task has been found to be faster and more stereotyped than previously thought (Kuchibhotla et al., 2019). We trained mice to respond to one tone for a water reward (S+) and withhold from responding to another (S-). We interleaved reinforced trials with those where reinforcement was absent (“probe”). Early in learning, animals discriminated between S+ and S- in probe but not reinforced trials. This unmasked a rapid acquisition phase of learning followed by a slower phase for reinforcement, termed ‘expression’. What role does cholinergic neuromodulation play in task acquisition? Here, we test the hypothesis that cholinergic neuromodulation provides a ‘teaching signal’ that drives primary auditory cortex (A1), and links stimuli with reinforcement. We exploit our behavioral approach and combine this with longitudinal two-photon calcium imaging of cholinergic axons in A1 during discrimination learning. We report both robust stimulus-evoked cholinergic activity to both S+ and S- and stable licking-related activity throughout learning. While this activity mildly habituates in a passive control, in behaving animals the S+ and S- stimulus-evoked activity is enhanced (S+: duration, S-: amplitude and duration) on the timescale of acquisition. Additionally, we test the hypothesis that cholinergic neuromodulation impacts the rate of task acquisition. We expressed ChR2 bilaterally in cholinergic neurons within the basal forebrain of ChAT-Cre mice and activated these neurons on both S+ and S– trials throughout learning. Test animals acquired the task faster than control groups. These results suggest that phasic bursts of acetylcholine, directly impact the rate of discrimination learning. In two additional cohorts, we paired the activation with either the S+ or the S– stimulus. In each case, we observed a refinement of the action associated with the paired stimulus: increased licking for S+ and suppression of licking for S–. Together, these results suggest that CBF activation differentially modulates contingency-related actions.

ID: 333

Abstracts (late breaking)

Keywords: mismatch negativity, pitch, location, intensity, p3a

Heterogeneous effects of inharmonicity on mismatch responses to different auditory features

Krzysztof Basiński[1], Tomasz Domżalski[1], Kornelia Renachowska[1], Zofia Mikołajczak[1], Alejandro Omar Blenkmann[2,3]

[1]Auditory Neuroscience Laboratory, Department of Psychology, Medical University of Gdańsk, Poland; [2]Department of Psychology, University of Oslo, Norway; [3]RITMO Centre for Interdisciplinary Studies in Rhythm, Time, and Motion, University of Oslo, Norway; k.basinski@gumed.edu.pl

Harmonicity is a fundamental acoustic feature that influences many important perceptual processes, for example pitch extraction, auditory memory or spoken language comprehension. Previous research suggests that inharmonicity captures attention and induces an auditory scene segregation process. However, since these results were obtained using pitch deviances in oddball experiments, it is unclear if they apply to other auditory features than pitch. In the present study, we used electroencephalography to measure the mismatch responses of the brain in response to a variety of auditory deviants in a multi-feature Optimum-1 paradigm (pitch, intensity, location, omission and harmonicity). We systematically manipulated the harmonicity of complex tones by introducing random jitters to frequency components above the fundamental. These patterns could either be the same within a block, or changing with each consecutive sound. We studied N = 29 healthy volunteers (22 female), aged M = 22.0 SD = 1.9 with a 64 channel active EEG system. Cluster-based permutations analysis revealed statistically significant differences in MMN amplitude between changing and harmonic conditions for pitch (estimate = +2.17 μV, $p = .0013$) and intensity (+1.87 μV, $p = .0061$), but not for location (+0.06 μV, $p = .9943$). This indicates that harmonicity influences pitch and intensity perception, but not location. Inharmonic deviant in a harmonic context produced different responses than a harmonic deviant in an inharmonic context. Omission responses are not modulated by harmonicity ($p > .05$). Overall, these results point to heterogeneous influence of inharmonicity on mechanisms of deviance detection depending on the auditory feature.

ID: 334
Abstracts (late breaking)
Keywords: cochlear implant, multisensory, hearing and speech development, deafness

Assessment of sensory integration in children with cochlear implants rehabilitated using the verbotonal method

Angela Nikolic Margan[1], Andro Kosec[2]
[1]SUVAG Polyclinic, Zagreb, Croatia; [2]University Hospital Centre Sisters of Charity, Zagreb, Croatia; lina.margan@gmail.com

When it comes to the senses and their impairments, there are two well-known theories: the deprivation theory and the compensation theory. According to the deprivation theory, the loss of one sense weakens the other senses, while according to the compensation theory the loss of one sense strengthens the remaining ones. In the rehabilitation of children with cochlear implants, the senses are an extremely important component that is difficult to examine precisely because of the complexity of the sensory systems. Sensory integration is a term that denotes the brain's ability to organize all stimuli it receives from the environment through the senses into a meaningful and functional whole. There are tests that have been designed specifically to assess sensory integration. The aim of this study is to assess sensory integration patterns in children with cochlear implants (CI) rehabilitated using the verbotonal method. Assessing sensory integration patterns in children with CIs will provide longitudinal insight into the maturation processes of sensory systems due to the implantation of CIs, and in the short term, it will enable the adaptation of rehabilitation procedures to the child's sensory specificities, if they exist. Three groups of children with congenital deafness that have been in the process of speech and hearing rehabilitation using the verbotonal method since the moment of getting a CI will undergo the EASI (Evaluation of Ayres Sensory Integration) testing. All of the hearing impaired participants are patients of the Cochlear Implant Centre at SUVAG Polyclinic in Zagreb, Croatia. In order to analyze the different periods of children's sensory development, the subjects will be divided into three groups that differ in chronological age and duration of rehabilitation. The first group will consist of children aged 3 to 4 years, the second group of children aged 6 to 7 years, and the third group of children aged 8 to 9 years. The participants in the groups will be paired by gender, cochlear laterality, and type of hearing impairment in such a way that the groups are homogeneous and comparable, and differ only in the duration of hearing rehabilitation. The fourth group will consist of children with normal hearing, who will form the control group that will be selected from a healthy kindergarten population with no recorded hearing impairment. The results will then be compared with previously validated EASI test norms and with a control group of children who do not have a record of hearing impairment. Since this is an ongoing doctoral research, preliminary results will be presented.

ID: 335
Abstracts (late breaking)
Keywords: statistical learning, TRF, EEG, cortical tracking, language

The neural representation of structured sound sequences following game-based learning

Nathaniel Zuk
Nottingham Trent University, United Kingdom; nathaniel.zuk@ntu.ac.uk

Statistical learning theories propose that humans acquire speech and music through long-term exposure by extracting patterns in what they hear, yet most studies use passive listening paradigms that neglect the interactive nature of real-world communication. Here, we created a novel, game-based paradigm to establish an active context that mimics real-world listening scenarios. Half of the participants were assigned to active, game-based training and learned to associate strings of modulated tone syllables with in-game ingredients when placing food orders in an alien language. Participants spent from 30 minutes up to one hour and 15 minutes playing through the game and learning six two- and three-syllable ingredient "words". The other half of the participants did passive training, where they listened to the sounds from a recorded game session for 45 minutes while playing a different game. All participants then performed a behavioural test of their knowledge of the sounds, followed by a passive listening session to continuous sequences of concatenated true words and non-words while recording EEG. Actively trained participants did better on the behavioral task than passively trained participants, demonstrating better retention of the in-game words. Using temporal response function (TRF) analysis, we found significant syllable tracking in the EEG responses to the sound sequences. But there was little evidence, in both training groups, of additional predictive power when the higher-level structure was included in the TRF model. These preliminary results imply that cortical tracking of higher-level expectations in structured, speech-like sound sequences requires long-term experience.

ID: 336
Abstracts (late breaking)
Keywords: Sound prediction, action binding, action kinetics

Attributes of sound outcome shape the kinetics of the actions that generate them

Roy Mukamel
Tel Aviv University, Israel; rmukamel@tau.ac.il

Motor actions are typically seen as means to manipulate the environment and achieve intended sensory outcomes. However, emerging evidence highlights a dynamic interplay, where anticipated sensory feedback actively shapes the kinetics of motor execution. For example, it has been shown that when associating a button press with auditory feedback, participants apply less force compared to when the same button is not associated with sound outcome.

In the current study, we further explore this interplay and examine how attributes of sound feedback interact with the force profile of the actions that generate them. We used a force-sensitive plate to measure the force profiles of right index finger taps while the contingency between finger-presses and auditory outcome (300ms, 1kHz pure tone) was manipulated. Healthy, young participants (N = 24) tapped to trigger an auditory sound of either high or low amplitude across experimental conditions (73 or 23 dB SPL, respectively). Results show an inverse relationship such that when outcome amplitude is known in advance, participants applied more force when the expected sound intensity was low (vs. high intensity sound). However, when sound amplitude could not be predicted in advance, force profiles were similar across the two sounds. When flipping the causal order between action and sound such that high/low amplitude sounds precede the action, no amplitude-dependent modulations were found (N = 24; different set of participants). The results of these studies show that attributes of associated sound outcome are embedded in the kinetics of the actions that generate them.

It is not uncommon, that sound feedback is associated with the final movement of a multi-movement action sequence. Therefore we also examined whether associated sound outcome affects kinetics of early movements within the sound generating movement sequence. Participants (N = 40) performed single or double taps which generated (or not) a sound immediately following the last tap in the sequence. Sound-dependent force modulations were seen also in the force profile of the first tap in the double-tap sequence. In other words, the applied force in the first (silent) tap was lower when the second tap in the sequence was associated with a sound compared to when it was not. Thus, the associated sound affected the kinetics of a temporally and causally separate earlier movement.

Taken together, our results demonstrate that sound attributes shape action kinetics in early and late stages of movement. We are now examining such interplays in bimanual actions.

ID: 337

Abstracts (late breaking)

Keywords: auditory-motor learning, N100, P50, surprisal, music

Auditory-Motor Surprisal Reveals Context-Sensitive Dynamics in a Novel Music Task

Cindy Haiqin Zhang[1], Giorgia Cantisani[1], Shihab Shamma[1,2]

[1]Department of Cognitive Studies, École Normale Supérieure, France; [2]Department of Electrical and Computer Engineering, University of Maryland College Park; cizhang21@gmail.com

Auditory-motor learning is important in learning how to produce and control sounds such as in speech and playing music. As in all sensorimotor learning, it is based on internal models of interactions between actions and their sensory consequences, which are fine-tuned by minimizing the errors ("surprisal") between the received sound and those predicted from the internal model. Here we apply the concept of surprisal to probe the neural dynamics of sensorimotor learning in the context of a keyboard-playing task in which participants played on a keyboard that changes between three key-to-pitch maps in random order and at short time intervals while their EEG was recorded during playing, listening, and miming (playing the keyboard without sound). We hypothesized that staying within the same key-to-pitch map engenders the smallest errors (surprisals), compared to switching between different maps. To test this, we extracted a neural signature of auditory-motor surprisals by comparing first keystroke after a map change (highly surprising events) to all other keystrokes (less surprising events). We found a signature of auditory-motor surprisal, the N100, which is not fully explained by pure auditory surprisal. Further analyses showed that this surprisal signature is found in the auditory response rather than the motor response. The magnitude of N100 is modulated by short-term context, with greater surprise after a long period without surprising events, showing that in addition to using short-term context to predict the auditory-motor mapping, the brain tracks environmental volatility in real time. Long-term musical training further enhances N100 amplitude. In contrast, P50 amplitude is modulated by 30-minutes of goal-directed training, and is apparent only in the auditory-motor mapping that was trained. This experiment presents a novel approach to studying auditory-motor surprisal using time-varying stimuli from the perspective of online surprisal. Our findings reveal that auditory-motor learning is fast, context-sensitive, and shaped by both short- and long-term experience.