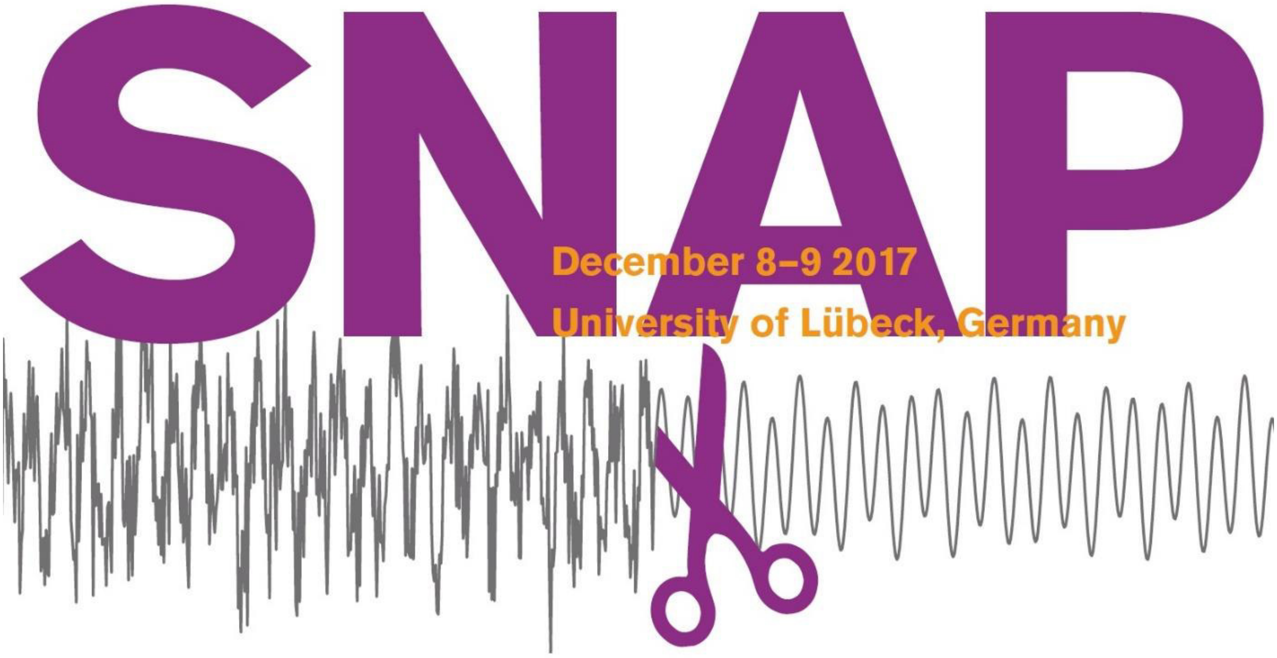
UNIVERSITÄT ZU LÜBECK
IM FOCUS DAS LEBEN



Signal and Noise along the Auditory Pathway

International workshop on current neuroscience and engineering perspectives on processing degraded speech and sound

Venue
Atlantic Hotel
Schmiedestraße 9–15
23552 Lübeck

Organized by the Obleser lab, auditorycognition.com

UNIVERSITÄT ZU LÜBECK

erc
European Research Council
Established by the European Commission

Oticon Foundation

sivantos
the hearing company

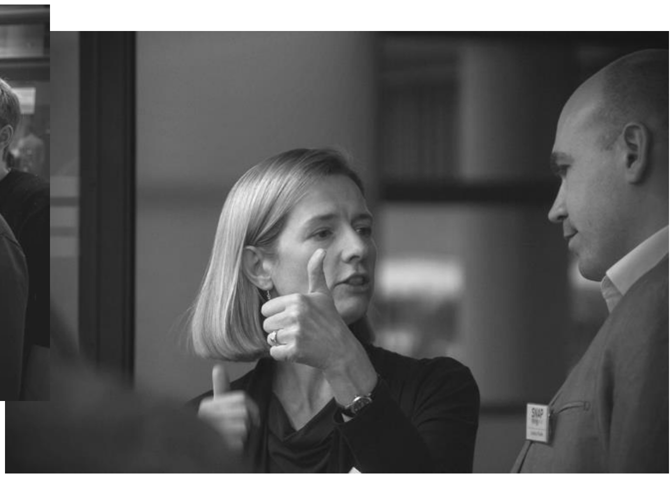
MES
FORSCHUNGSSYSTEME GMBH

BRAIN PRODUCTS
Solutions for neurophysiological research

audifon
hearing systems

Table of Contents

Impressions from SNAP 2013 in Leipzig



Program

Friday, December 8, 2017

08:00 – 09:00 am	Workshop Registration
09:00 – 09:10 am	Opening remarks
09:10 – 10:00 am	Edmund Lalor *"The effects of attention and visual input on noninvasive electrophysiological indices of natural speech processing at different hierarchical levels"*
10:00 – 10:50 am	Sarah Verhulst *"Model-based design of subcortical EEG methods to quantify different aspect of sensorineural hearing loss"*
10:50 – 11:10 am	Coffee Break
11:10 – 12:00 pm	Christoph Kayser *"Audio-visual speech encoding – contributions of local representations and functional connectivity"*
12:00 – 12:30 pm	Discussion (Impulses by Sonja Kotz)
12:30 – 02:00 pm	Lunch
02:00 – 02:50 pm	Tom Francart *"EEG in response to running speech: applications in diagnostics and noise suppression"*
02:50 – 03:40 pm	Nima Mesgarani *"Robust speech processing in human auditory cortex"*
03:40 – 04:00 pm	Coffee Break
04:00 – 06:00 pm	Poster Session
07:00 pm	Speakers Dinner at Buddha Bowl Schmiedestraße 26
09:00 pm	SNAP Party at "Gang No 56" Marlesgrube 58

Saturday, December 9, 2017

09:30 – 10:20 am	Jennifer Bizley *"How does auditory cortex maintain information content in complex listening conditions?"*
10:20 – 11:10 am	Samira Anderson *"Age related declines in sensory-cognitive processing: Clinical implications"*
11:10 – 11:30 am	Coffee Break
11:30 – 12:20 pm	Maria Chait *"The role of the LC-NE system in tracking the changing statistics of rapid sound sequences"*
12:20 – 12:50 pm	Discussions (Impulses by Thomas Münte)
12:50 – 02:00 pm	Lunch
02:00 – 02:50 pm	Josh McDermott *"New Models of Auditory Cortical Computation"*
02:50 – 03:30 pm	Round-table / Wrap-up

General Information

Baggage Check

There is limited storing space available for luggage. Please contact the registration desk.

Certificate of Attendance

Certificates of attendance can be provided upon request. If you need a certificate of attendance, please contact Franziska Scharata (franziska.scharata@uni-luebeck.de).

Food Service

Complementary food and beverage service is available to all registered attendees.
Light refreshments will be available throughout the days along with several dedicated coffee breaks in between session. A lunch buffet will be provided on both days of the workshop.

Internet Access

Free WiFi is available at the meeting space. Please see the registration desk for log-in details.

Poster Session

The poster session will take place on Friday afternoon in Conference Room 1. Poster presenters are encouraged to put up their poster in the morning to give attendees the chance to browse through the poster session areas during coffee breaks as well.
The poster board dimensions are
H:140cm x W:114cm (portrait).

Program

Our program booklet including the abstracts of talks and poster presentations is available for download at:
http://auditorycognition.com/snap2017-booklet

Social Events

Please join us on Friday evening, starting at 9 pm for the legendary SNAP Party taking place at *Gang No.56* in Marlesgrube 58 (see map on the back of the booklet). Gang No. 56 is located only a short walk away from the workshop hotel and offers a truly authentic Lübeck experience and a great selection of drinks (no host bar).

Social Media

Follow the SNAP discussion on Twitter!

Tag meeting-related tweets with #SNAP2017

There is also a prize for the best picture taken at or around the SNAP workshop, posted on Twitter using the hashtag #SNAP2017shot

Speakers

We will provide a laptop enabled to present both PowerPoint and Keynote presentation formats but speakers are also welcome to use their own computers. Please get in touch with our on-site technical staff before your session to test your setup.

Registration

The registration desk will open on Friday the 8th at 8:00 am and at 8:30 am on Saturday the 9th.

Speakers and Talks

Edmund Lalor

University of Rochester / Trinity College Dublin

Edmund Lalor is an Associate Professor of Biomedical Engineering and of Neuroscience at the University of Rochester, US. His background is in electrical and biomedical engineering and he received a PhD in Biomedical Engineering from the University College Dublin. Having completed post-doctoral training in the Cognitive Neurophysiology Laboratory at the Nathan Kline Institute for Psychiatric Research in New York, and at the Institute of Neuroscience and at the Centre for Bioengineering at Trinity College Dublin, he became an Ussher Lecturer at Trinity College Dublin.

His work seeks to explore quantitative modelling approaches to the analysis of sensory electrophysiology in humans with a particular focus on the processing of natural stimuli such as speech, music, and video, as well as the integration of visual and auditory information in speech comprehension.

The effects of attention and visual input on noninvasive electrophysiological indices of natural speech processing at different hierarchical levels

Friday, December 8, 2017, 9:10-10:00 am

Chair: Lea-Maria Schmitt

How the human brain extracts meaning from the dynamic patterns of sound that constitute speech remains poorly understood, especially in natural environments. In this talk I will outline efforts over the last few years to derive noninvasive indices of natural speech processing. I will discuss how these indices are affected by attention and visual input and how attentional selection and multisensory integration can be "decoded" from EEG data. I will outline work showing that EEG and MEG are sensitive not just to the low-level acoustic properties of speech, but also to higher-level linguistic aspects of this most important of signals. This will include demonstrating that these signals reflect processing at the level of phonetic features. And, it will also include evidence that EEG is exquisitely sensitivity to the semantic processing of natural, running speech in a way that is very strongly affected by attention and intelligibility.

Sarah Verhulst

Ghent University

Sarah Verhulst is an Associate Professor of Hearing Technology at Ghent University. She has a background in electrical and acoustical engineering and received her PhD from the Technical University of Denmark on the topic of cochlear mechanics and otoacoustic emissions. After post-doctoral training in computational auditory neuroscience at Boston and Harvard University, she became Assistant Professor in physiology and modelling of auditory perception at Oldenburg University.

In her research, she adopts an interdisciplinary approach to study sound encoding in the normal and impaired auditory system. Her work combines physiological such as EEG and otoacoustic emissions with computational modelling of the auditory periphery to study how different aspects of peripheral hearing loss interaction and affect sound perception.

Model-based design of subcortical EEG methods to quantify different aspect of sensorineural hearing loss

Friday, 8 December, 2017, 10:00 – 10:50 am

Chair: Jens Kreitewolf

Even though animal physiology studies have demonstrated that cochlear synaptopathy affects the amplitude of auditory brainstem responses (ABRs) and envelope-following responses (EFRs), the sensitivity of these subcortical EEG metrics in diagnosing synaptopathy in humans is unclear. Confounding factors include head size, gender and other forms of peripheral hearing loss (e.g., outer-hair-cell (OHC) loss) that may co-exist and affect subcortical EEG amplitudes as well. In this work, we use a computational model of the auditory periphery that simulates the broadband and level-dependent features of human ABR and EFRs to optimize subcortical EEG stimulation paradigms for synaptopathy. We simulated how OHC loss or dysfunction impacts the sensitivity of ABR and EFR metrics of synaptopathy and show that: (i) a combination of the slope between the EFR fundamental and 1st harmonic together with the ABR/EFR amplitude ratio, or (ii) the ABR growth ratio that takes into account both the amplitude and latency growth of ABRs, can map listeners onto a OHC loss vs synaptopathy degree scale. We compared the simulations to recordings from listeners with normal and sloping high-frequency audiograms to evaluate the practical use of the proposed relative subcortical EEG metrics.

Christoph Kayser

Universität Bielefeld

Christoph Kayser is a Professor of Cognitive Neuroscience at the University of Bielefeld, Germany. He has a background in mathematics, theoretical physics, as well as molecular, clinical and system neuroscience and received his doctorate degree in Neurobiology from the ETH Zürich. Prior to his current appointment in Bielefeld, he held a position as Professor of Integrative Neuroscience at the University of Glasgow, UK.

His research focuses on multisensory perception and multisensory integration in the brain and combines methods ranging from single unit recordings to functional imaging such as M/EEG, psychophysics, and computational approaches.

Audio-visual speech encoding – contributions of local representations and functional connectivity

Friday, 8 December, 2017, 11:10 – 12:00 am

Chair: Michael Plöchl

While many brain regions have been implicated in audio-visual speech integration, the neural correlates of behavioural benefits remain unclear. We studied these correlates in two MEG studies. In one we quantified speech representations in temporally entrained brain activity. During high acoustic SNR speech encoding was strong in temporal and inferior frontal cortex, while during low SNR encoding emerged in premotor and superior frontal cortex. This was accompanied by changes in directed connectivity along the ventral stream and the auditory-premotor axis, some of which were directly predictive of the behavioural benefit. In another study we investigated visual benefits for the representations of single words using multi-variate decoding. This revealed a network of temporal, parietal and inferior frontal regions, in which the read-out of word identity was predictive of perception. These findings highlight the need to consider both local representations and functional connectivity when trying to elucidate the neural underpinnings of speech perception, and highlight how the ventral and dorsal pathways facilitate speech comprehension in challenging environments.

Tom Francart

Katholieke Universiteit Leuven

Tom Francart is a Research Professor in the Research Group Experimental Oto-rhino-laryngology (ExpORL) at KU Leuven, Belgium. Following his M.A. studies in electrotechnical engineering, he received a PhD from KU Leuven for his work on the perception of binaural localization cues with combined electric and acoustic hearing. He completed post-doctoral training both at ExpORL at Ku Leuven, as well as at the Bionics Institute in Melbourne, Australia.

His work follows a multidisciplinary approach that links electrical engineering with audiology and neuroscience. His current research focuses on the development of individualised and self-adapting sound processing for cochlear implant users.

EEG in response to running speech: applications in diagnostics and noise suppression

Friday, 8 December, 2017, 2:00 – 2:50 pm

Chair: Lorenz Fiedler

The envelope of a natural speech signal can be decoded from the EEG. The correlation between the decoded signal and the original envelope yields a measure of neural entrainment to the envelope. As the speech envelope is a primary cue for speech intelligibility, we hypothesis that its neural representation is a prerequisite for speech intelligibility. It can therefore be used to predict speech intelligibility if most other influencing factors can be accounted for. We investigated the interrelations between speech intelligibility, neural entrainment of the envelope, listening effort, and attention. By carefully selecting speech materials, procedures and signal processing methods, we were able to find a correlation between speech intelligibility and neural entrainment of the speech envelope. This has applications in objective audiometry and closed-loop auditory prostheses that adapt their function to an individual user.

Nima Mesgarani

Columbia University

Nima Mesgarani is an assistant professor of Electrical Engineering at Columbia University. He received his Ph.D. from University of Maryland where he worked on neuromorphic speech technologies and neurophysiology of mammalian auditory cortex. He was a postdoctoral scholar in Center for Language and Speech Processing at Johns Hopkins University, and the neurosurgery department of University of California San Francisco before joining Columbia in fall 2013. He was named a Pew Scholar for Innovative Biomedical Research in 2015, and received the National Science Foundation Early Career Award in 2016.

His research focuses on information processing of acoustic signals at the intersection of engineering and neuroscience. Using an interdisciplinary research approach, he aims to bridge the gap between these two disciplines by reverse engineering the signal information processing in the brain, which in turn inspires novel approaches to emulate human abilities in machines.

Robust speech processing in human auditory cortex

Friday, December 8, 2017, 2:50-3:40 pm

Chair: Mohsen Alavash

The brain empowers humans with remarkable abilities to navigate their acoustic environment in highly degraded conditions. This seemingly trivial task for normal hearing listeners is extremely challenging for individuals with auditory pathway disorders, and has proven very difficult to model and implement algorithmically in machines. In this talk, I will present the result of an interdisciplinary research effort where invasive and non-invasive neural recordings from human auditory cortex are used to determine the representational and computational properties of robust speech processing in the human brain. These findings show that speech processing in the auditory cortex is dynamic and adaptive. These intrinsic properties allow a listener to filter out irrelevant sound sources, resulting in a reliable and robust means of communication. Furthermore, incorporating the functional properties of neural mechanisms in speech processing models greatly impact the current models of speech perception and at the same time, lead to human-like automatic speech processing technologies.

Jennifer Bizley

University College London

Jennifer Bizley is a Sir Henry Dale Research fellow and Reader in Auditory Neuroscience at the UCL Ear Institute in London, UK. Previously she was a PhD student in Neuroscience at the University of Oxford, working with Profs Andy King and Jan Schnupp, where she was also a post-doc and Royal Society Dorothy Hodgkin Research Fellow.

Her main research interest is in how the brain makes sense of the barrage of sensory information available to it and includes research topics such as auditory scene analysis, perceptual invariance and auditory-visual integration. A particular focus of her work is on the neural mechanisms of listening in noisy and complex –yet everyday–listening conditions. This includes understanding how visual information and attention enable us to listen effectively.

How does auditory cortex maintain information content in complex listening conditions?

Saturday, 9 December, 2017, 9:30 – 10:20 am

Chair: Jonas Obleser

In this talk I will discuss a series of recent experiments in our lab in which we explore how neural representations are maintained across perceptual constancy, as well as in the presence of background noise and competing sounds. We obtained data from the auditory cortex of ferrets while they discriminated the identity of a synthetic vowel across variation in pitch, level and location. We describe a multiplexed and multivariate code that enables a neural representation that is robust to identity-preserving manipulations while still sensitive to changes in other feature dimensions. We expand these studies to determine the impact of background noise with different temporal characteristics on behavioural performance and neural encoding. Finally, I will present data showing that auditory-visual temporal coherence can determine the response of single neurons in auditory cortex to a sound mixture, biasing neurons towards representing the temporally coherent sound, and enhancing the representation of both binding and non-binding acoustic features.

Samira Anderson

University of Maryland

Samira Anderson is an Assistant Professor in the Department of Hearing and Speech Sciences at the University of Maryland, US. After practising as a clinical audiologist for 26 years, she received a PhD in Communication Sciences and Disorders from Northwestern University before joining the University of Maryland as a faculty member.

Her work focuses on the effects of development, aging, and hearing loss on central auditory processing, using electrophysiology, in particular auditory brainstem measures, as her primary assessment tool. She also aims to define neuroplasticity in response to different forms of treatment, such as hearing aid amplification, cochlear implantation, and auditory training.

Age related declines in sensory-cognitive processing: Clinical implications

Saturday, 9 December, 2017, 10:20 – 11:10 am

Chair: Sarah Tune

Older adults often experience difficulty understanding speech in challenging listening situations. These difficulties may arise in part from deficits in neural speech processing that arise from changes associated with aging and hearing loss. To compensate for impaired perception, older adults rely on cognitive resources to a greater extent than do younger adults, despite the fact that cognitive function decreases with age. If declines in neural speech processing and cognitive function are not taken into account, it may be challenging to predict the amount of benefit an older adult will receive from amplification or other forms of intervention. This presentation will discuss the interacting effects of aging, hearing loss, and cognition on neural processing and speech perception. Clinical implications will be discussed in terms of hearing loss management through the use of devices and training.

Maria Chait

University College London

Maria Chait is a Professor of Auditory Cognitive Neuroscience at the UCL Ear Institute in London, UK. Her background is in Computer Science, Economics, and East Asian Studies and she received a PhD in Neuroscience and Cognitive Science from the University of Maryland College Park. In 2007, she moved to UCL as a Marie Curie research fellow, following a short post-doc at Ecole normale supérieure, Paris.

Her research focuses on how listeners use sounds in order to learn about, and interact with their surroundings. Using psychophysical methods, eye tacking and functional brain imaging, she explores how representation that are useful for behaviour arise from sensory input and dissociate automatic, stimulus-driven processes from those that are affected by the perceptual state, task, and goals of the listener.

The role of the LC-NE system in tracking the changing statistics of rapid sound sequences

Saturday, 9 December, 2017, 11:30 – 12:20 pm

Chair: Malte Wöstmann

Sensitivity to patterns is fundamental to sensory processing and a major component of the influential ‘predictive coding’ theory of brain function. Supported by growing experimental evidence, the ‘predictive coding’ framework suggests that perception is driven by a mechanism of inference, based on an internal model of the signal source. However, a key element of this theory - the process through which the brain acquires this model, and its neural underpinnings – remains poorly understood. Our experiments focus on this missing link. I will present the results of ongoing experimentation in my lab which aims to reveal how human listeners discover patterns and statistical regularities in rapid sound sequences. I will specifically focus on work using pupillometry to tap the role of the LC-NE system in the process of tracking the statistical structure of the environment.

Josh McDermott

MIT Massachusetts Institute of Technology

Josh McDermott is an Assistant Professor in the Department of Brain and Cognitive Sciences at the Massachusetts Institute of Technology, US. Josh has a background in Computational Neuroscience and received his PhD in Brain and Cognitive Science from MIT. His post-doctoral training in psychoacoustics at the University of Minnesota marked his transition into auditory research, having received his early training in vision research.

In his research, which operates at the interaction of psychology, neuroscience, and engineering, he seeks to understand how humans derive information from sound, to improve treatments for those whose hearing is impaired, and to enable the design of machine systems that mirror human abilities to interpret sound.

New Models of Auditory Cortical Computation

Saturday, 9 December, 2017, 2:00 – 2:50 pm

Chair: Leonhard Waschke

The ability to derive information from sound is enabled by a cascade of neuronal processing stages that transform the sound waveform entering the ear into cortical representations that are presumed to make behaviorally important sound properties explicit. Although much is known about the peripheral processing of sound, the auditory cortex is less understood, particularly in humans, with little consensus even about its coarse-scale organization. This talk will describe our recent efforts to develop and test models of auditory cortical computation, to delineate function within auditory cortex, and to understand the role of the cortex in robust sound recognition.

Discussants

Sonja Kotz

Maastricht University

Sonja Kotz is a Professor and Chair of Neuropsychology and Psychopharmacology in the Department of Psychology and Neuroscience at Maastricht University, NL. Her early background is in German Philology, English studies and political science. She then received a Master's and PhD degree in Experimental Psychology/Neuropsychology from Tufts University, US. Before joining Maastricht University, Sonja led an independent Max Planck-Minerva research group at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig, and held the position of Chair in Cognitive and Affective Neurosciences at the University of Manchester.

Her work covers a wide range of topics, ranging from social communication, time perception and predictive coding tin auditory, visual, and multimodal perceptions to work on learning and neuroplasticity, as well as cognitive and affective control.

Thomas Münte

Universität zu Lübeck

Thomas Münte is a Professor of Neurology and Head of the Clinic for Neurology at the University of Lübeck, Germany. Following his medical education, during which he spent time as a graduate student in the Department of Neuroscience as the University of California, San Diego, he received his medical doctorate degree from the Hannover Medial School. Much of his early work focused on electrophysiological correlates of language processing. Prior to joining the University of Lübeck in 2010, Thomas held a position as Professor and Chair of Neuropsychology and Head of the Clinic of Neurology at the University of Magdeburg.

At the University of Lübeck, he leads a research group on Cognitive Neurology that employs a wide range of experimental methods including neuropsychological testing, functional and structural brain imaging, electrophysiology, intracranial recordings, transcranial magnetic stimulation, as well as genetic testing to help understand the basis of neurological and neuropsychiatric diseases such as Parkinson's disease, amyothrophic lateral sclerosis, obsessive compulsive disorder, or dyslexia.

Poster Presentations

P.1 L-dopa modulates signal variability and the functional connectome in the listening brain

Mohsen Alavash[1,2], Sung-Joo Lim[1,2], Christiane. M. Thiel[3], Bernhard Sehm[2], Lorenz Deserno[2], Jonas Obleser[1,2]
[1] Institute of Psychology, University of Lübeck, Germany
[2] Max Planck Research Group "Auditory Cognition", Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[3] Biological Psychology Lab, Department of Psychology, European Medical School, Carl von Ossietzky Universität Oldenburg, Germany

Dopamine underlies important aspects of cognition, and has been suggested to boost cognitive performance. However, how dopamine modulates the large-scale cortical dynamics during cognitive performance has remained elusive. Using functional MRI during a working memory task in healthy young human listeners (N=22), we investigated the effect of levodopa (L-dopa) on two aspects of cortical dynamics, blood oxygen-level-dependent (BOLD) signal variability and the functional connectome of large-scale cortical networks. We here show that enhanced dopaminergic signaling modulates the two potentially interrelated aspects of large-scale cortical dynamics during cognitive performance, and the degree of these modulations is able to explain inter-individual differences in L-dopa-induced behavioral benefits. Relative to placebo, L-dopa increased BOLD signal variability in task-relevant temporal, inferior frontal, parietal and cingulate regions. On the connectome level, however, L-dopa diminished functional integration across temporal and cingulo-opercular regions. This hypo-integration was expressed as a reduction in network efficiency and modularity in more than tow third of the participants and to different degrees, and co-occurred with a relative hyper-connectivity in paracentral lobule and precuneus, as well as posterior putamen. Both, L-dopa-induced BOLD signal variability modulation and functional connectome modulations proved predictive of an individual's L-dopa-induced gain in behavioral performance, namely response speed and perceptual sensitivity. Lastly, L-dopa-induced modulations of BOLD signal variability were correlated with L-dopa-induced modulation of nodal connectivity and network efficiency. Our findings underline the role of dopamine in maintaining the dynamic range of, and communication between, cortical systems, and how these relate to inter-individually differing benefits from dopamine during cognitive performance.

P.2 Auditory attention decoding in reverberant and noisy conditions

Ali Aroudi[1], Simon Doclo[1,2]
[1] Department of Medical Physics and Acoustics, University of Oldenburg, Germany
[2] Cluster of Excellence Hearing4all, University of Oldenburg, Germany

During the last decades significant advances in multi-microphone acoustic signal processing algorithms have been achieved to improve speech intelligibility for hearing-impaired listeners. Nevertheless, enhancing speech in complex acoustic environments, particularly in multi-talker scenarios, is still a challenging problem since most algorithms rely on predefined assumptions about the target speaker to be enhanced. Recently, a least-squares auditory attention decoding (AAD) method has been proposed for identifying the attended speaker from single-trial EEG recordings. While extensive research efforts have focused on auditory attention decoding in dichotic and anechoic conditions, the feasibility of AAD in reverberant and noisy conditions still remains unclear. In addition, the considered AAD method requires the clean speech signals of both the attended and the unattended speaker to be available as reference signals. However, in practice only the binaural signals consisting of a reverberant mixture of both speakers and background noise are available. In this study, we explore the feasibility of decoding auditory attention in different acoustic conditions (anechoic, reverberant, noisy, and reverberant-noisy). We show that using the unprocessed binaural signals as reference signals for decoding auditory attention is feasible. In addition, we show that for

obtaining a good decoding performance the joint suppression of reverberation, background noise and interference as undesired acoustic components is of great importance.

P.3 Behavioral and Neural Representations of Auditory Looming Bias Demonstrate the Effect of Degraded Spectral Cues on Spatial Perception

Robert Baumgartner[1,2], Darrin K. Reed2, Brigitta Tóth[3], Virginia Best4, Piotr Majdak[1], H. Steven Colburn[2], Barbara Shinn-Cunningham[2]
[1] Acoustics Research Institute, Austrian Academy of Sciences, Vienna, 1040, Austria;
[2] Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA;
[3] Institute of Cognitive Neuroscience and Psychology, Hungarian Academy of Sciences, Budapest, 1117, Hungary;
[4] Department of Speech, Language & Hearing Sciences, Boston University, Boston, MA 02215, USA

Sound sources are naturally perceived as externalized auditory objects located outside the head. When listening via headphones or hearing-assistive devices, sounds are often heard inside the head, presumably because individual spectral cues become inconsistent with normal experience. The present study evaluated the consequence of changes in spectral cue salience on sound externalization. Listeners judged differences in externalization while the salience of spectral cues was varied by modifying the spectral contrast of listener-specific head-related transfer functions (HRTFs). Behavioral results showed strong individual differences, but, on average, smaller spectral contrasts were judged to be less external. In simultaneous electroencephalographic recordings, spectral changes caused event-related potentials characterized by an early fronto-central negativity (N1) around 120 ms and a later fronto-central positivity (P2) around 220 ms. N1 and P2 magnitudes were larger the more the spectral contrast changed. Moreover, decreasing spectral contrasts lead to larger P2 magnitudes and higher behavioral response consistency than symmetrically increasing spectral contrasts. This is consistent with previous reports on the perceptual salience of looming auditory objects, usually attributed to sounds increasing versus decreasing in overall sound intensity. The auditory looming bias arguably reflects increased alertness for approaching objects and thus confirms that reduced salience of spectral cues markedly degraded sound externalization.

P.4 Reconstruction of noise-vocoded speech from primary auditory cortex during perceptual pop-out

Alexander J. Billig[1], Ariane E. Rhone[2], Phillip E. Gander[2], Kirill V. Nourski[2], Conor J. Wild[1], Matthew A. Howard III[2], Ingrid S. Johnsrude[1];
[1]Brain and Mind Institute, The University of Western Ontario,
[2]Human Brain Research Laboratory, Department ofNeurosurgery, The University of Iowa

Speech comprehension under adverse listening conditions can be enhanced by semantic context and visual information (Miller et al., J Exp Psychol 41:329-35, 1951; Sumby & Pollack, J Acoust Soc Am 26:212-5, 1954). For noise-vocoded sentences lacking fine spectrotemporal detail, matching text facilitates perceptual "pop-out" such that the impoverished signal is reported as less noisy than when heard alongside meaningless letter strings (Wild et al., Neuroimage 60:1490-502, 2012). Wild et al. (2012) found an elevated BOLD response in auditory cortex associated with this effect, suggesting that early cortical stages of auditory processing can be modulated by non-auditory information in support of speech comprehension. Here, we investigated whether corresponding increased activity can be observed in high gamma power recorded intracranially from primary auditory cortex in patients undergoing monitoring for epilepsy, and if so whether this reflects differences in overall attention across conditions, or changes in the tuning properties of neural populations in response to visual information. In each trial, subjects were presented with an auditory

sentence (clear or 3-channel noise-vocoded), then written text, followed by a repeat of the auditory stimulus. In the clear speech case, the text matched the auditory sentence; in the noise-vocoded cases text could be matching or mismatching. To engage attention, subjects were asked to detect occasional capital letters in the text and warned that memory for the sentences would be tested at the end of the experiment. Changes in event-related high gamma (75-150 Hz) power were computed for each auditory presentation, relative to a pre-stimulus baseline. First, responses to noise-vocoded speech were averaged over contacts in postero-medial Heschl's gyrus most likely to be recording from primary auditory cortex. For at least one subject, a significant cross-over interaction between text type and presentation number reflected an increase in high gamma power from the first to the second presentation when the text was matching, compared to a decrease when text was mismatching. The design rules out such effects being driven by acoustic differences or mere repetition. We next demonstrated that the high gamma power timecourses at 4-8 electrode contacts per subject in primary auditory cortex are sufficient to reconstruct clear speech above chance (mean correlation between actual and reconstructed spectrograms, r=.22). Next we will test whether increased primary auditory cortical activity following matching (compared to mismatching) visual information reflects the retuning of neural populations to better match anticipated auditory input. To do so we will compare how well spectrograms of noise-vocoded sentences (and of the clear sentences from which they are derived) can be reconstructed from neural activity under the different visual conditions. Better reconstruction accuracy following matching than mismatching text would be consistent with rapid anticipatory retuning in human primary auditory cortex in response to visual information.

P.5 Mapping the latent space of human auditory functional specialization with binary sparse coding

Moritz Boos[1], Jörg Lücke[2], Jochem W. Rieger[1]
[1] Applied Neurocognitive Psychology Lab, University of Oldenburg, Oldenburg, Germany
[2] Machine Learning Group, University of Oldenburg, Oldenburg, Germany

In humans, brain activity can be predicted from a given stimulus representation combining fMRI with voxelwise encoding models (Naselaris et al., 2011). Yet, the choice of stimulus representation can limit the interpretability of the encoding model. In the visual system (Güçlü and van Gerven, 2014) showed that features learned with sparse coding better predict BOLD fMRI than hand-crafted Gabor wavelets. In the auditory domain, an efficient coding of natural sounds accounts for tuning properties of auditory nerve fibers (Lewicki, 2002). Our aim here, was to go beyond prediction, by testing if encoding models based on auditory features learned with an unsupervised approach from natural auditory stimuli can inform us about principles of cortical functional organization in the auditory domain. For 15 participants from an open 7-Tesla fMRI dataset (Hanke et al., 2014), we predict BOLD activity elicited by an auditory movie. We learn a representation of the Mel-frequency spectrogram of the auditory movie using sparse coding with binary latents (BSC) (Lücke and Eggert, 2010; Henniges et al., 2010). We decompose the encoding model into latent dimensions, using principal component analysis. This reveals three dimensions that are shared between participants. The high similarity between participants allows us to build an encoding model that generalizes across different brains, predicting data for a left-out participant. The first dimension of across-participants encoding model is located mostly in Heschl's gyrus and correlates highly with stimulus energy. The other two dimensions are located more laterally and ventrally (than the first dimension) along the superior temporal sulcus. The second and third dimension are sensitive to the auditory complexity of stimuli, they span dimensions between simpler and more complex stimuli, while the first dimension shows no such sensitivity. In conclusion, unsupervised learning, together with a generalizing encoding model that is shared across participants, reveals general principles of the functional anatomical organization of human temporal cortices.

P.6 The Music-In-Noise Task: a tool for dissecting complex auditory perception

Coffey, Emily B.J., Arseneau-Bruneau, I., Lim, Ah Ryung, and Zatorre, Robert .J.; Montreal Neurological Institute, McGill University, Montréal, Québec, Canada H3A 2B4; Laboratory for Brain, Music and Sound Research (BRAMS), Montréal, Québec, Canada; Centre for Research on Brain, Language and Music (CRBLM), Montréal, Québec, Canada H3G 2A8

The ability to understand speech in the presence of competing sound is relevant both to neuroscience and clinical applications. Recent research suggests that hearing-in-noise (HIN) problems are solved using combinations of sub-skills that are applied according to task demand and information availability. We have developed the Music-In-Noise Task (MINT) as a flexible tool for addressing research questions such as the relative contributions of HIN sub-skills, inter-individual differences in their use, and their neural correlates. The MINT uses a match-mismatch trial design and includes four conditions (Pitch, Rhythm, Spatial, and Visual) in which subjects first hear a short instrumental musical excerpt embedded in "multi-music babble", followed by either a matching or slightly altered repetition of the musical excerpt presented in silence. In the fifth condition, subjects hear the music as a target first, the memory of which helps to predict incoming information. Data from samples of young control subjects who differ in their exposure to musical training (N=70) and the relationships between MINT conditions and additional behavioural measures associated with HIN will be presented (i.e. auditory working memory, speech-in-noise, and fine pitch discrimination thresholds). A customizable version of the MINT is made available for use by the scientific community.

P.7 Encoding of mid-level speech features in Magnetoencephalography responses

Christoph Daube, Robin A. A. Ince, Joachim Gross
Institute for Neuroscience and Psychology, University of Glasgow, Glasgow, UK

Studying the neuronal responses to naturalistic speech stimuli offers the perspective of retracing increasingly abstract transformations of the auditory input as it is processed along the auditory pathway. One computational approach to explore hypotheses about such transformations in Magneto- or Electroencephalography (MEEG) data relies on the estimation of multivariate temporal response functions (mTRFs). With this class of regularized linear models, previous studies have demonstrated the relevance of spectrally resolved stimulus energy and linguistically derived phonemic features for the prediction of MEEG signals. This makes it interesting to investigate what intermediate acoustic features afford the listener such pre-lexical abstraction. To address this question we recorded the MEG of young healthy human participants (currently N=3) while they listened to a narrative of 1 hour duration. We performed source localization using LCMV beamformer and identified regions of interest based on retest-reliable responses to a repeated speech stimulus. In these regions surrounding bilateral Auditory Cortices we estimated mTRFs from bandpass-filtered signals (1-15 Hz) in a 5-fold nested cross-validation framework and compared the predictive power of a number of different feature spaces: Plain stimulus energy, cochleagrams, time-series of modulation power spectra (MPS), a set of automated voice feature measurements and a set of linguistic articulatory features. In all participants, we replicated the finding of improved prediction performance when articulatory features were added to cochleagrams. We achieved a comparable increase in prediction performance when we instead added MPS features to cochleagrams. We consistently obtained best prediction performances when we used all feature spaces in combination. An examination of the variance of learned weights of the four feature spaces in the full model revealed sensitivity to onset detection in cochleagrams, highlighted regions of the MPS known to be crucial for speech comprehension, exposed contributions of measures of spectral tilt and stressed articulatory features describing manner- and place of articulation. Our results

demonstrate that mTRFs combining various feature spaces can be estimated reliably for individual grid points from source-localised MEG data. This allowed us to extend previous acoustic models with further mid-level auditory features. Taken together, this is a promising next step to fill the gap between cochlear and linguistic representations of speech stimuli using temporally highly resolved and non-invasive neuroimaging.

P.8 Neural responses to concurrent speech reflect the emergence of an SNR-invariant representation of the attended talker

Lorenz Fiedler, Malte Wöstmann, Sophie Herbst & Jonas Obleser
University of Lübeck, Lübeck, Germany

Listening conditions in real-life scenarios are rarely constant but vary unpredictably. While the focus of auditory attention can be detected from EEG, varying acoustical conditions may involve different neural mechanisms responsible for amplification of the attended signal (here, a talker) and/or suppression of the ignored noise (here, another talker). Here, we investigated the effect of varying signal-to-noise ratio (SNR) on the neural response to concurrent speech. Eighteen participants attended to one of two simultaneously presented, spatially non-segregated talkers, while we measured their EEG. The SNR was varied by stochastic variation in the sound intensity of both the signal (attended) and noise (ignored) such that changes of SNR occured unpredictably. Using a forward encoding-model approach, neural responses to the attended and ignored talker were estimated. We investigated the influence of SNR (i.e. bottom-up) and attentional (i.e. top-down) manipulations on both amplitude and phase of the neural response functions. Across conditions, responses to the attended talker showed a P1-N1-P2 complex in all subjects, whereas the N1-P2 components to the ignored talker were largely suppressed. First, the magnitude of the P1 component solely depended on the SNR and tracked the louder talker. However, the P1 response to the ignored talker was phase-delayed, which suggests that the attended signal becomes prioritized relatively early. Second, the N1 to the attended (vs. the ignored) talker was larger in all SNR-levels, but this difference increased in magnitude with more favorable SNRs. Lastly, the P2 component clearly contrasted the attended from the ignored talker, but this contrast was constant across SNRs. Interestingly, a late negative, N2-like deflection occurred only in responses to the ignored talker and only under the worst SNR. This additional component localized to parietal regions (whereas all other effects localized to superior and middle temporal regions). This is likely a signature of fronto-parietal attention network activity actively suppressing the ignored talker in a top-down manner. This chain of the P1-N1-P2 complex as revealed by a forward-encoding model approach and their sensitivity to temporally unpredictable changes in SNR indicate that the neural representation of the attended talker gets successively isolated from the distractor. This was also reflected in enhanced classification accuracy in the P2 interval across conditions when utilizing a temporal generalization method to identify the attended talker. Our findings give insight into the spatio-temporal unfolding of the neural cortical response, and how an attentional set can help overcome acoustic challenges along the auditory pathways.

P.9 Decoding complex sounds from activity patterns

Dupont Typhaine [1,2,3], Michalski Nicolas [1,2,3], Petit Christine [1,2,3,5,6], Gourevitch Boris [1,2,3,4]
[1]Unité de Génétique et Physiologie de l'Audition, Institut Pasteur, Paris, France,
[2]UMRS 1120, Institut National de la Santé et de la Recherche Médicale (INSERM),
[3]Sorbonne Universités, UPMC Université Paris 06,
[4]CNRS, France,
[5]Syndrome de Usher et Autres Atteintes Rétino-Cochléaires, Institut de la Vision,
[6]Collège de France, Paris.

Neurons from the auditory system have long been characterized through their SpectroTemporal Receptive Field (STRF), which is basically a linear approximation of the spectrogram having evoked a spike. This method, while useful for neuron features assessment, has been shown to perform poorly in predicting neural activity from acoustically complex sounds (coding) or reconstructing sounds from spike trains (decoding). This may be due to the specific neural code in the auditory system, and especially in the auditory cortex, which is based on very precise spike timing. Insofar as this is the case, coincident spiking and temporal patterns of spikes support most of the information about acoustic cues, and these features are largely ignored by STRFs. In this preliminary study, we present a simple and intuitive jackknife-like decoding method to predict stimulus information from multiple spike trains, here applied to auditory processing. We show that temporal patterns of spikes allow for the accurate reconstruction of complex acoustic cues based on the activity of small numbers of neurons. We then compare reconstruction performance of acoustic stimulus between neurons from the inferior colliculus, the thalamus and the primary auditory cortex of mice. This very generic decoding method can also be reversed to perform a completely non-parametric prediction of evoked neural firing rate.

P.10 Do EEG-based artificial neural networks detect the attentional focus of a listener?

Simon Grosnick, Lorenz Fiedler & Jonas Obleser
Department of Psychology, University of Lübeck, Germany

A listener's focus of attention in a multi-speaker environment (cocktail party problem) can be detected from multi-channel EEG data. Does the application of artificial neural networks lead to any benefits in terms of higher classification accuracy, compared to linear approaches? In a concurrent listening task, two audio books were presented simultaneously to eighteen participants. They were asked to attend one of two simultaneously presented talkers while a 64-channel EEG was recorded. A forward encoding model in the form of an single-layer neural network was trained to predict a multi-channel EEG-response based on the speech stimuli. Comparing the predicted EEG-response to the empirically measured one provides information about the subjects focus of attention. Classification accuracies that are similar to studies that apply linear methods were achieved, confirming that the auditory focus of a person can be encoded from multi-channel EEG. Typical characteristics of auditory evoked potentials were revealed in the network's response. In addition, a non-linear relationship in the networks data processing was found. Although the classification accuracies indicate no significant advantage for the application of neural networks, the investment might lead to a better understanding of non-linear relationships between an auditory stimulus and the neural responses it evokes.

P.11 Implicit temporal predictability enhances auditory pitch-discrimination sensitivity

Sophie K. Herbst, Michael Plöchl, Anne Herrmann & Jonas Obleser
Department of Psychology, University of Lübeck, Germany

The human brain extracts temporal contingencies from the environment, to orient attention to future events and prepare behavioural responses. It is a matter of debate, however, how automatic such temporal prediction is. A putative neural mechanism for temporal prediction has been proposed via phase alignment of slow neural oscillations, such that the best phase for stimulus processing coincides with the predicted time point of target occurrence (Schröder & Lakatos, 2009). It is, however, unclear whether phase alignment is contingent on neural entrainment to rhythmic stimulation, or whether it also applies to single temporally predictive intervals. Here, we studied whether and how human listeners exploit strictly implicit temporal contingencies in an auditory pitch discrimination task: We presented tone pairs embedded in lowpass-filtered white noise, and asked participants to compare the pitch of a probe to a preceding standard. Critically, the standard pitch was either 550 Hz or 950 Hz, with both standards occurring randomly across trials. Unbeknownst to participants, one of the standard pitches was deterministically predictive of the time of occurrence of the probe, while the other standard pitch had no such predictive value (signaling only a uniform probability of probe occurrence; counterbalanced between participants). Both conditions were presented in one stream, separated by variable inter-stimulus intervals, such that there was no dominant stimulus rhythm throughout. We behaviorally assessed perceptual sensitivity to pitch deviation via the slope of the psychometric function (N = 51). Sensitivity increased when the onset of the target tone was predictable in time. Thus, participants extracted the implicit temporal contingencies present in the stimuli and utilized them to perform a sensory task, for which time itself was not a relevant dimension. In an EEG data set (N=24, also included in the behavioural sample), we examined the neural mechanisms of processing temporal predictions in a non-rhythmic paradigm. Standard tones were treated differentially with respect to whether they functioned as a temporal cue: Standard tones which initiated temporal predictions evoked a more negative N1 component than non-predictive standards. Furthermore, delta power was enhanced during temporally predictive foreperiods, confirming that slow neural oscillations play a crucial role in temporally predictive processing. Delta ITC was not significantly stronger during predictive foreperiods, indicating that such a mechanism might require entrainment over more than one trial (see also Herbst & Obleser 2017). However, temporal predictions evoked a more efficient delta phase reset by the standard tone, which in turn was related to pitch discrimination sensitivity.In sum, the data show that human listeners automatically exploit implicit temporal contingencies present in sensory input. Our EEG data provide evidence for an underlying neural mechanism of a more adaptive phase reset evoked by temporally predictive cues.

P.12 The effect of distraction on auditory change detection: an ageing study

Mathilde de Kerangal and Maria Chait
Ear Institute, University College London, London, UK

Elderly people show deficits in a variety of cognitive tasks. For example, auditory decline can be manifested by difficulties in segregating auditory objects in a crowded scene or difficulties in inhibiting irrelevant information. This study investigates the effect of distraction using an auditory change detection paradigm in two age groups (young: 18-35 years old, old: over 60 years old). Artificial scenes composed of multiple sequences of tones playing simultaneously were used to model a complex environment. The change that the participants were asked to detect was of the appearance or disappearance of a component in the scene. A distracting sound was added

concurrently with this change. It is hypothesized that the addition of the distracting sound would lead to a drop in performance for both groups, but that this would affect more the older listeners than the younger ones. The study presents a correlation analysis that examined the relation between the change detection performances and measures reflecting auditory and attentional decline (audiogram, speech in noise task, sustained attention task).

P.13 Talker familiarity interferes with working memory processes in an irrelevant-speech task

Jens Kreitewolf, Malte Wöstmann, Jonas Obleser[1]
Department of Psychology, University of Lübeck, Lübeck, Germany

Familiarity with a talker's voice improves speech comprehension, not only when listeners attend to a familiar talker's speech but also when they ignore it in the context of a cocktail party. Here, we investigated the effect of talker familiarity on speech comprehension in an irrelevant-speech task. We assessed listeners' working memory for the serial order of spoken digits when a task-irrelevant, distracting sentence was produced by either a familiar or an unfamiliar talker. The listeners were N = 66 undergraduate psychology students who had attended an introductory statistics course. Critically, each student had received classroom teaching by one of two lecturers, whose voices served as familiar and unfamiliar task-irrelevant talkers in the present experiment. We thus were able to manipulate talker familiarity within subject. The two task-irrelevant talkers were presented in alternating blocks. In about 17 % of the trials, the task-irrelevant sentence was omitted (silence trials). The results showed that talker familiarity affected the performance in the irrelevant-speech task. Interestingly, listeners tended to recall less digits in the familiar talker blocks compared to the unfamiliar talker blocks. This familiarity deficit was evident for task-irrelevant sentences (but not for silence trials) and in the first half of the experiment (but not in the second half). Our findings suggest that talker familiarity cannot only improve comprehension of attended speech, but it can also deteriorate speech comprehension through interference with working memory processes.

P.14 Effects of voice continuity and stimulus rate on auditory working memory

Sung-Joo Lim[1,2,] *Tyler K. Perrachione*[1]*, and Barbara G. Shinn-Cunningham*[1,2]
[1]*Department of Speech and Hearing Sciences, Boston University*
[2]*Departments of Cognitive and Neural Systems and Biomedical Engineering, Boston University*

Speech perception can be cognitively demanding when incoming speech is highly variable. The auditory system prefers to adapt to speech from a single talker, but this speaker adaptation process requires both cognitive resources and time. However, the facilitatory effect of speaker adaptation has been mostly investigated under tasks requiring immediate recognition of speech. Thus, it is unclear whether the facilitatory effect of speaker adaptation is also present for speech information being maintained in auditory working memory, and how this effect changes when listeners have more time to process speech from multiple talkers. Using a delayed recall of digit span task, we investigated whether time for memory rehearsal of speech mitigates interference of processing multiple talkers, and whether this potential reduction of interference changes as a function of speech rate at encoding. On each trial, listeners encoded a seven-digit sequence and recalled it after a 5-s delay. Sequences were spoken by single or multiple different talkers and were presented at three different rates (0, 200, or 500 ms inter-digit intervals). We found that recall was less accurate and slower for sequences spoken by multiple talkers than by a single talker. This interference of processing multiple talkers was most evident at faster rates of digit sequence presentation. However, slower rates of sequence encoding generally improved listeners' memory recall overall. These results suggest that the speaker adaptation effect persists in auditory working memory items, especially

when listeners are not provided with time to encode and adapt to speech from multiple talkers. These behavioral results provide the foundation to further investigate electrophysiological correlates of how speech variability influences auditory encoding and working memory processes.

P.15 Brain mechanisms involved in the use of visual speech to compensate for acoustically degraded speech

Heidi S. Økland, Lucy MacGregor, Saskia Helbling, Helen Blank & Matt Davis
University of Camebridge, Camebridge, UK

Seeing the person we talk to can compensate for loss of acoustic information in adverse listening conditions, leading to better speech comprehension both in healthy people and those with hearing loss (Erber, 1975; Summerfield, 1992). Our project focuses on this visual enhancement of speech perception in normal-hearing people. We ask two main questions: (1) What are the neural mechanisms underlying this visual enhancement of speech perception with degraded acoustic input? (2) How does visual enhancement of speech perception vary between individuals in behaviour and neural recordings? Participants in our study watch videos of a speaker saying everyday English sentences where the audio has been noise-vocoded to produce varying degrees of intelligibility. Audiovisual (AV), auditory-only (AO, neutral face) and visual-only (VO, no sound) trials are presented in pseudorandom order, and participants verbally report the words they could perceive after each video. We record participants' responses and calculate the percentage of words that matched the target sentence (word report). In our first behavioural study ($n = 11$), we found a clear audiovisual benefit for word report ($F(1,50) = 119.776$, $p < 0.001$) starting at around 30% for the most severely degraded condition and peaking at around 60% for intermediate levels of degradation, before dropping off when AO word reports approached ceiling ($F(5,50) = 33.720$, $p < 0.001$). There was a large amount of individual variation in AV word report at low and intermediate acoustic clarity which, interestingly, was not present for AO word report. This suggested that the degree to which participants were able to use the visual speech available to them was highly variable. The word report scores for VO comprehension (i.e. speechreading) supported this interpretation, as being a good speechreader correlated positively with performance for the most severely degraded AV sentences ($r = 0.888$, $p < 0.01$). We are currently running a combined MEG/EEG study of to look at the neural processing that underlies visual enhancement of speech perception and how this varies between individuals. Phase coherence between quasi-rhythmic components of both acoustic and visual speech is emerging as a potential mechanism by which speech perception might occur in the brain (e.g. Peelle & Davis, 2012; Park, Kayser, Thut & Gross, 2016). Previous studies have shown that neural oscillations show significant phase coherence with both the acoustic speech envelope and its visual counterpart, lip aperture. Making use of coherence as an index for neural speech processing, we will investigate (a) whether neural coherence with the speech envelope and/or lip aperture changes with differences in acoustic clarity and the presence or absence of visual speech and, (b) how individual differences in the use of visual information are reflected in coherence between visual speech and cortical areas involved (audio)visual speech processing. The poster will present preliminary results from ongoing analyses of the MEG and behavioural data.

P.16 Are visual and auditory detection performance driven by a supramodal attentional rhythm?

Michael Plöchl[1], Sabine Kastner[2], Ian Fiebelkorn[2] & Jonas Obleser[1]
[1] Department of Psychology, University of Lübeck, Germany
[2] Princeton Neuroscience Institute, Princeton University, Princeton, NJ

Recent research indicates that visual attention, and thus the likelihood of detecting faint visual target stimuli, is not constant over time, but fluctuates rhythmically at a rate of ~8–12 Hz. When two objects are attended simultaneously, attention samples these objects at about half of the original rate (4-6 Hz) and in counter-phase. However, whether a similar attentional rhythm also exists in the auditory modality is still a matter of debate. Here we adapt and extend an established paradigm to investigate fluctuations of both visual and auditory attention. We explore, whether attentional rhythms are modality-specific or potentially governed by a supramodal mechanism.
Participants performed an audio-visual detection task, in which they had to detect near-threshold target stimuli (contrast changes, or beeps within red noise). This was performed either in a unimodal condition, with purely auditory or purely visual stimulation, or in a bimodal condition, where targets could occur equally likely in both modalities. We measured detection performance (N=24) and neural oscillatory activity (EEG, N=12) under the hypothesis that target detection depends on the time (0.3–1.2 s) and location (same/opposite) at which the target appears relative to a salient auditory/visual event ("flash event" to reset attention towards one object/side). As hypothesized detection performance fluctuated most prominently at frequencies between 3 and 8 Hz. We compared phase differences between attentional oscillations associated with "same-side" and "opposite-side" detection performance, respectively. In the unimodal conditions the largest phase angle differences in the visual modality occurred at ~4 Hz and in the auditory modality at ~4 Hz and ~8 Hz, indicating that at those frequencies attention towards the "same" and the "opposite" location fluctuates almost in counter phase. In the bimodal conditions the 4 Hz peaks were shifted towards 3 Hz, although in the auditory condition it did not exceed the 95% confidence level. Furthermore, we investigated EEG patterns in response to visual and auditory flash-events. A direct contrast between trials with left and right flash-events confirmed that the respective visual responses are lateralized at 2-5 Hz at occipital electrodes. Moreover, in the presence of visual flash-events we found significant phase opposition ($p = 0.004$) in those frequencies where we also observed the strongest power increase and lateralization (3-8 Hz), but this time peaking at parietal rather than occipital electrode sites. This suggests that phase opposition is not brought about by a low-level response to the stimulus per se but possibly relates to a higher-level process such as an attentional reset. We found no lateralization or counter-phasic activity in response to purely auditory flash-events. Finally, preliminary analyses point towards the possibility that the observed counter-phasic theta activity (~3–8 Hz) in response to left and right flash events is also reflected in modulations of alpha power (9–12 Hz). Overall our preliminary results suggest that attentional rhythms (1) also exist in the auditory modality, (2) are endogenous (i.e., not necessarily driven by entrainment to external rhythms), (3) alternate between objects/locations, and (4) are potentially governed by a supramodal oscillatory mechanism.

P.17 Cortical tracking of auditory and audiovisual speech in hearing-impaired and normal-hearing old Listeners

Sebastian Puschmann[1], Mareike Daeglau[2], Maren Stropahl[2], Stefanie Rosemann[2], Christiane Thiel[2], Stefan Debener[2]
[1]Montreal Neurological Institute, McGill University, Montreal, QC, Canada
[2]Department of Psychology, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany

Hearing-impaired subjects are commonly thought to rely on visual cues (e.g., lip movements) to support speech perception in challenging listening situations (Hallam & Corney, 2014). It is however open whether hearing-impaired listeners make more use of additional visual information than normal-hearing listeners (Tye-Murray et al, 2007). To investigate this question, we here compared behavioural performance and EEG markers of speech envelope tracking in moderately hearingimpaired (HI) and normal-hearing (NH) old subjects when listening to blocks of naturalistic auditory and audiovisual speech in noise. Visual-only speech blocks controlled for a changed processing of visual input in the HI group. Participants (HI: N=20, mean age: 64 ± 7 years; NH: N=20, mean age: 62 ± 6 years) listened to randomly distributed blocks of auditory, visual, and audiovisual speech and had to respond via button press to rare target words (30 blocks of 30 sec per condition). The level of the background noise was adjusted individually for each listener to the 80% speech reception threshold. During the experiment, EEG was recorded from 64 equidistant scalp electrodes. The EEG data was preprocessed in EEGlab. The mTRF toolbox (Crosse et al., 2016) was used for the speech envelope reconstruction. The behavioural data show significant main effects of listening condition (AV>A, $p<.001$) and group (NH>HI, $p=.011$) on correct target detection. No interaction between both factors was observed ($p=.071$). Complementing the behavioral data, first results of the EEG data analysis show a significant main effect of condition on the speech envelope reconstruction (AV > A > V, $p<.001$). There was no significant main effect of group ($p=.905$) and no group by listening condition interaction ($p=.870$). In summary, our preliminary results suggest that visual information facilitates speech processing in challenging listening situations, as indicated by a superior behavioral performance in AV compared with A only trials and improved cortical tracking of AV speech streams. Hearing-impaired listeners however seem to not specifically benefit from visual cues in this rather difficult task.

P.18 A musical cocktail party: Understanding polyphonic music through the mismatch negativity (MMN)

David Quiroga, Niels C. Hansen, Andreas Højlund, Elvira Brattico, Peter Vuust
Aarhus University, Aarhus, Denmark

The mismatch negativity (MMN) is a brain response elicited by deviants in a series of repetitive sounds. It reflects the perception of change in sound features and reliably measures auditory memory and predictive processes. The MMN is a valuable tool for the study of music perception. However, most experimental designs use simple tone patterns as stimuli, failing to represent the complexity of everyday music. Our goal was to develop a new MMN paradigm using more real-sounding stimuli. Concretely, we wanted to use the MMN to assess the perception of full non-repetitive melodies when presented alone and when embedded in a polyphonic context, namely, two-part music. A previous design based on the Alberti bass served both as a comparison and as the second voice in the two-part stimuli. We used MEG to record nonmusicians' responses to four types of deviants (mistuning, intensity, timbre and slide), while they watched a silent film under four conditions: listening to bass only ("bass"), listening to melody only ("melody"), listening to bass in the pitch range of the melody ("bass high"), and listening to bass and melody together ("together"). We found MMNs for all deviants for the "melody" condition. However, for mistunings and slide deviants, MMNs were reduced compared to the "bass high" condition. This is likely due to the higher

pitch complexity of the melodies. Moreover, we found a reduction of MMNs for most features in the two-part excerpts, which can be explained by masking in the cochlea or competition for neural resources. Interestingly, this reduction did not hold for mistunings and slide deviants in the melody, probably due to a combination of vertical sensory dissonance and the high voice superiority effect. Our results indicate that it is possible to study auditory perception with the MMN using more complex and real-sounding stimuli, such as non-repetitive melodies. However, the complexity of the context plays a crucial role, especially in polyphonic music. Further research is needed to reduce masking/neural competition effects in order to study the perception of simultaneous musical streams.

P.19 Age-related hearing-loss modulates audio-visual speech processing

Stephanie Rosemann[1,2], Christiane M. Thiel[1,2]*
[1] Biological Psychology, Department of Psychology, European Medical School, Carl-von-Ossietzky Universität Oldenburg, Oldenburg, Germany
[2] Cluster of Excellence "Hearing4all", Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany

Previous research provides compelling evidence for a cross-modal reorganization in the auditory cortex following auditory sensory deprivation, leading to increased neural responses to visual input. Recently, it was shown that these changes may already occur after a moderate hearing-impairment. By using functional magnetic resonance imaging, we investigated the influence of mild to moderate hearing-impairment on audio-visual speech processing. Twenty participants with mild to moderate hearing-loss who do not yet wear a hearing-aid and nineteen normal-hearing age-matched participants performed an audio-visual speech detection task in which they had to indicate which word was included in the previous sentence. Sentences were presented either audio-visually congruent, audio-visually incongruent, only visually or only auditorily. Furthermore, the participants performed a working memory task, an audio-visual integration task (McGurk illusion), and a questionnaire assessing hearing effort. Analysis of the behavioral data indicate significantly higher hearing effort and stronger audio-visual integration for the hearing-impaired subjects. Unimodal auditory and visual stimuli as well as incongruent audio-visual stimuli elicited an increased frontal activation pattern in hearing-impaired individuals compared to normal-hearing subjects, possibly related to an increased effort to perform the task. Hearing-loss modulated both the audio-visual integration strength and brain activation in frontal areas, showing stronger integration and higher brain activation with increasing hearing-loss. Our data provide evidence for hearing-loss related changes in audio visual speech processing as well as BOLD response amplitudes in frontal brain areas.

P.20 Spectrally resolved neural responses to concurrent speech in a listener with otosclerosis

Sarah Sentis, Lorenz Fiedler & Jonas Obleser
Department of Psychology, University of Lübeck, Lübeck, Germany

In the past decade, it has been shown that spectrally resolved neural responses (i.e. spectro-temporal response functions, STRF) to continuously presented auditory stimuli can be extracted from electrophysiological recordings. STRFs might be a suitable tool to fit hearing prosthesis. How individual characteristics of the audiogram and the processing of concurrent auditory stimuli are reflected in STRFs is unknown. Here we ask whether STRFs of a single participant suffering from the conductive hearing loss otosclerosis give insight to the selective neural processing of concurrent talkers. We recorded EEG while the participant was asked to attend to one of two simultaneously presented talkers (female and male talker) for the listening conditions with and without hearing aids.

Equivalent EEG recordings of 18 normal hearing subjects served as control data. To obtain an estimation of the STRFs, a linear forward encoding model was trained using ridge regression. In the frequency ranges affected by the hearing loss, significant attention-related differences between the two listening conditions (with and without hearing aids) were observed. With hearing aids, a stronger positive neural response to the attended talker was observed at a time lag around 200 ms in the frequency range affected by otosclerosis. This indicates, that the attended talker yields a stronger neural representation due to the hearing aids. Contrastingly, a stronger response to the ignored talker was observed without hearing aids in the frequency range not affected by otosclerosis, which indicates that hearing aids help to suppress the ignored talker. Our findings show that the effect of otosclerosis (and its treatment with hearing aids) on the selective neural processing of concurrent auditory stimuli is reflected in STRFs. Thus, STRFs might be suitable to individually fit hearing aids in order to improve selective listening.

P.21 Is tonal working memory a reliable predictor of speech understanding in hearing impaired listeners?

Maja Serman, Maren Schuelke, Kaja Kallisch
Sivantos GmbH, Erlangen, Germany

The role of cognition on speech understanding in hearing impaired listeners has been widely investigated. The picture is complex, since speech comprehension tests may differ with respect to their cognitive demands. The relationship between the cognitive demands of such tests and real-life situations is also unclear (Heinrich et al 2015). Accordingly, we were interested in investigating the relationship between speech understanding and working memory for basic sound attributes such as pitch, loudness and directionality. Here, we report on two studies investigating tonal working memory and performance in different speech in noise scenarios. For this purpose, we used a *memory for tones* task consisting of auditory rhythmic sequences of low frequency tones (a same/different judgment of a standard and a target tone, separated by a series of six intervening distracting tones). The results show that listeners who performed well in the tonal working memory task had significantly better speech-in-noise performance than listeners whose tonal working memory test performance was bad. This was true both for a cognitively demanding, spatial speech in noise setup (N = 30, $p < .000$) and for an easier, closed sentence setup with speech and noise coming from the front speaker (N = 20, $p < .05$).

P.22 Auditory scene ambiguity and the impact of prior context: the case of timbre

Kai Siedenburg
Dept. of Medical Physics and Acoustics
University of Oldenburg, Oldenburg, Germany

Recent research has revealed profound effects of prior acoustic context on pitch shift judgements of ambiguous tone pairs (Chambers et al., 2017, Nat. Comm.). The study mainly utilized so-called "Shepard tones" with octave-spaced partials and a fixed bell-shaped spectral envelope. These tones are long known to exhibit ambiguous pitch percepts for half-octave shifts (Shepard, 1964, JASA). Generalizing from Shepard tones, we here propose the concept of "Shepard spectra". We define a Shepard spectrum as a harmonic tone complex with a two-level spectral envelope: a global bellshaped spectral envelope (fixed across tones) in conjunction with a more fine-grained local spectral envelope (varied between tones). Notably, shifts in the local spectral envelope do not affect the perceived pitch of tones but only the perceived timbral brightness. The goal of this study is to examine whether shifts of the local spectral envelope of Shepard spectra and the resulting changes in brightness yield effects of ambiguity and auditory context analogous to the ones found for pitch.

Two experiments (modeled after Chambers et al., 2017, Exps. 1&2) are currently conducted to test the hypotheses that i) there is an ambiguity of brightness perception for half-octave shifts of local spectral envelope in Shepard spectra; and ii) there can be strong effects of auditory context on brightness shift judgments, analogous to the ones found for pitch shifts. Beyond highlighting important similarities between the processing of pitch and timbral brightness, a confirmation of these hypotheses would extend current knowledge of the principles by which auditory scene ambiguity emerges and the ways in which it is resolved by prior context.

P.23 Selective attention in a spectro-temporally complex auditory scene: a ferret cocktail party

Sollini, Joseph
University College London, London, UK

A key function of the auditory system is the segregation of competing sounds. An example of this is the Cocktail Party phenomenon, where a listener is able to attend to one talker in the presence of one or more competing talkers. Animal behavioural models have been employed to investigate the cocktail party phenomenon, though these models generally focus on conspecific vocalizations rather than human speech. Here we tested the feasibility of using an animal model in a selective attention task using human speech, while making neural recordings from auditory cortex. We first established an animal model of the cocktail party phenomena. Ferrets (n = 5) were trained to respond to a target word embedded within a stream of distractor words, all performed significantly above chance (p<0.05, bootstrap, chance = ~26%, average = ~63% correct). Animals (n=2, thus far) were then trained to selectively attend to one of two competing streams and again respond to the target embedded within the attended stream. Animals performed above chance in the presence of babble speech (~47% correct) and another talker (~45% correct). Chronic recording electrodes were implanted bilaterally in auditory cortex and neural responses were measured during behavioural testing. As would be expected the stream onset responses were strongest in the contralateral hemisphere and were larger in situations where there was no simultaneous distractor stream. Preliminary results demonstrate the feasibility of this behavioural model potentially allowing closer inspection of the neural mechanism that underpins the cocktail party phenomena.

P.24 Sequence predictability is associated with enhanced error responses

Rosy Southwell1, Karl Friston[2], Maria Chait[1]
[1]Ear Institute, University College London
[2]Wellcome Trust Centre for Neuroimaging, University College London

How are responses to surprising events affected by the preceding context? In this series of behavioural and EEG experiments we measure responses to surprising 'deviant' sounds in regular or random sequences. We ask whether sequence predictability modulates the response to deviants, hypothesising increased evoked responses and better behavioural detection of deviants in a regular sequence. This would be consistent with predictive coding accounts, in which prediction errors are weighted according to the confidence, or precision, of the predictions [1]. Stimuli were rapid sequences of tone-pips, whose frequencies varied in a regular (REG) or random (RAND) pattern. REG sequences were created by repeating 10 frequencies in certain pattern; RAND patterns were created by shuffling the entire resulting sequence. REG and RAND patterns were overall matched for frequency content, and were controlled to contain no adjacent repeating tones. Deviants (in 50% of trials) were single tones higher or lower than the other frequencies in each stimulus. Deviants were equally divergent in frequency relative to both REG and RAND. Our paradigm is free from confounds due to adaptation of frequency-selective units, which is expected to be similar in both REG and

RAND. Listeners were faster and more accurate at detecting deviants in REG than RAND. A separate EEG experiment in naïve participants demonstrated sensitivity to regularly repeating patterns, manifest as a sustained increase in the evoked response to REG compared to RAND - consistent with MEG and fMRI results [2,3]. This is interpreted as reflecting increased weighting of the neural responses to regular stimuli, in order to focus processing resources on stable aspects of the environment. Importantly, over and above this difference, the EEG response evoked by deviants was larger in REG sequences. These findings demonstrate the sensitivity of automatic deviance detection to the level of regularity in the preceding context, consistent with predictive coding.

P.25 The cortical representation of sounds with speech-like modulation rates tested with multi-dimensional scaling

Xiangbin Teng1, David Poeppel1,2
1Max Planck Institute, Frankfurt, Germany
2Dept. of Psychology, New York University, New York, NY

Neural oscillations in auditory cortex can entrain to sounds over a wide range of modulations, although there appear to be preferential rates. Two such discrete regimes have been established and investigated, a low-frequency range (< 8 Hz) and a high-frequency range (> ~ 30 Hz). These regimes arguably reflect specialized temporal coding at two clearly distinct timescales in the auditory system. While cortical oscillations do not appear to be robustly entrained by sounds with intermediate modulation rates (8 – 30 Hz), such sounds are perceivable and still give people a clear sense of temporal structure. One particular question presents itself: how does the auditory system track and code the sounds with mid-range modulation rates? To address this question, we use a match-to-sample task in the context of MEG recording. Participants listen to sounds at five different modulation rates: 4-7 Hz ('theta' sound), 8-12 Hz ('alpha' sound), 13-20 Hz ('beta1' sound), 21-30 Hz ('beta2' sound), and 31-45 Hz ('gamma' sound). While undergoing MEG recording, participants executed a 3-IFC match to sample study with modulation rate as the critical variable. MEG analyses of phase locking as well as classification analyses illuminate the differential encoding – temporal dynamics of sounds can be only read out from theta and gamma bands but not others. Source-space analysis further shows that prominent entrainment of theta and gamma bands can be localized on primary auditory cortex and its surrounding areas. This result lends a strong support to a discrete multi-scale coding scheme on the cortical level of auditory processing (Poeppel, 2003).

P.26 Speech sound discrimination in noise: effects of adaptation on brain and behavior

Stephen M. Town, Tara Etherington, Katherine C. Wood, Jennifer K. Bizley
University College London, London, UK

The temporal context in which sounds arrive at the ear determines our experience an auditory scene. Here we ask whether adaptation to background noise could improve discrimination of degraded speech sounds and how adaptation affected auditory cortical processing. We trained subjects (14 humans; 6 ferrets) to discriminate artificial vowel sounds in a two-alternative forced choice task, and tested subject's discrimination in broadband noise that degraded sounds with a varying SNR (-30 to -10 dB). We compared two conditions in which noise was presented either during stimulus presentation only, or continuously in the background so that subjects could adapt to listening conditions. We hypothesised that a common temporal onset would make the noise harder to separate from the target sounds and result in elevated thresholds. Consistent with this, when discriminating targets in the presence of matched-SNR continuous or restricted both humans and ferrets performed better when vowels were presented in the continuous condition. Two trained ferrets were implanted with cooling loops to reversibly inactivate auditory cortex during task performance. Bilateral cooling did not impair vowel discrimination in silence but did impact performance in noise. While both noise conditions were affected, performance during cooling remained better when sounds were presented in continuous rather than restricted noise. Thus our results suggest a causal role for AC in the analysis of degraded sounds but that adaptation to continuous noise originates outside AC. To understand the mechanisms underlying behavior, trained ferrets were implanted with microelectrodes to record neural activity in auditory cortex during behavior (3 ferrets; 455 units in restricted noise, 285 units in continuous noise). Analysis of multi-unit activity revealed that temporally restricted noise reduced the proportion of units discriminating sound identity compared to when sounds were presented in silence, whereas a greater proportion of units discriminated sound identity in continuous noise than in silence. Across the population of recorded units, continuous noise suppressed baseline firing so that the target stimulus elicited similar responses to sounds as in silence. In contrast, restricted noise led to distorted onset activity which masked the response to the target vowel. Our findings show that the opportunity to adapt to noise improved behavioral sound discrimination and supressed cortical responses to the noise, allowing the cortical representation of the speech sounds to be maintained.

P.27 Auditory attention and predictive processing co-modulate speech comprehension in middle-aged adults

Sarah Tune, Malte Wöstmann & Jonas Obleser
University of Lübeck, Lübeck, Germany

In real-life communication, speech comprehension requires the dynamic engagement of a complex set of perceptual, executive control and prediction processes. This challenge becomes exacerbated by the gradual declines in sensory acuity and cognitive functioning that are typically associated with healthy aging starting in middle adulthood. Here, we present results from a study focused on healthy middle-aged human adults (40-70y) who performed a novel dichotic listening task. The paradigm called for adaptive control of cognitive strategies by varying the degree to which auditory spatial attention and predictive processing support comprehension. Participants were presented with two competing, dichotically presented speech streams uttered by the same female speaker. Participants were probed on the last word in one of the two streams. Crucially, auditory presentation was preceded by two visual cues. First, a spatial-attention cue either indicated the to-be-probed side, thus invoking selective attention, or it did not provide any information about the to-be-probed side, thus invoking divided attention. The second cue specified a general or a specific semantic category for the target word (and was valid for both utterances). This semantic cue therefore facilitated semantic and sensory prediction of the upcoming input. Behavioral results (n=29) show a general increase in performance for informative compared to uninformative cues. Participants responded faster in selective attention trials and following a specific semantic cue. Accuracy was co-modulated by the joint effect of both cues, as reflected by a benefit from specific (vs. general) semantic cues but only under selective auditory attention. Moreover, reliance on the spatial-attention cue varied with age: Older adults performed better under selective attention and worse under divided attention than younger adults. Analysis of electroencephalography (EEG) data (n=16) revealed a lateralization of 8–12Hz alpha power during spatial attention cue presentation, but also and even more pronounced during the dichotic speech streams in selective-attention but not in divided-attention trials. Specificity of the semantic cue on the other hand modulated oscillatory power in the beta frequency band (15–30Hz), with a decrease in power for specific cues. In sum, our results provide evidence for the interplay of attentional control and predictive processes in difficult listening situations. Crucially, providing two distinct types of cues prompted changes in behavioural performance correlated with qualitatively different neural signatures, and highlights changes in cognitive strategies with age.

P.28 The role of timing in automatic processing of speech at sub-lexical and lexical levels

Anita E. Wagner & Deniz Başkent
University of Groningen, University Medical Center Groningen, Department of Otorhinolaryngology
University of Groningen, Research School of Behavioral and Cognitive Neurosciences

Speech comprehension is automatic because the perceptual system enables instantaneous progression of information through various levels of analysis, from auditory object formation up to the integration of meaning within a conversation. How vital the timing of this progression is, becomes apparent when understanding speech becomes effortful, as it is for many profoundly deaf listeners, who perceive speech through a cochlear implant (CI). We study individual differences amongst CI users to investigate how the sensitivity to rhythmic structure in speech contributes to the automatic and effortless processing of speech via a CI. For individual listeners we created profiles of speech processing fluency based on cortical and behavioral measures that capture speech processing at various stages. Twelve experienced CI users, and twelve age-matched normal-hearing (NH) controls are compared in a set of measures: (1) cortical processing, measured as corticoacoustic coherence in EEG recordings when listening to continuous speech; (2) the time course of lexical access captured in gaze fixations in a visual world paradigm while resolving lexical ambiguities; (3) the mental effort involved while processing speech, as measured in pupil dilation. NH listeners show cortico-acoustic coherence in the delta (>4Hz) and theta (4-8 Hz) ranges. They also consistently rely on rhythm while resolving ambiguities, and show targeted and short-lived increase in mental effort during lexical access. For the CI users, individual variation in the mental effort involved in processing speech was related to individual differences in reliance on rhythm to resolve lexical ambiguities. Also, while all CI users showed increased cortico-acoustic coherence in the theta range, only CI users whose time-course of lexical access was similar to NH listeners showed coherence also in the delta range. This confirms the functional role of rhythmic structure in speech processing, but suggest also that extraction of this structure is more an active than automatic process.

P.29 Attention-modulated tracking of 1/f stimulus characteristics in the human EEG

Leonhard Waschke[1,2], Thomas Donoghue[2], Jonas Obleser[1] & Bradley Voytek[2]
[1] Dept. of Psychology, Univ. of Lübeck, Lübeck, Germany;
[2] Dept. of Cognitive Sci., Univ. of California San Diego, La Jolla, CA

The statistics of both human electrophysiology data and many natural sounds follow 1/f distributions. Recent works has proposed the 1/f slope of human electroencephalogram (EEG) activity spectra to track the balance of excitatory and inhibitory neural activity, which in turn manifests as a marker of interpersonal traits and behaviourally relevant states. It is unclear, however, to which degree such "tilt" of the electrophysiological spectrum can be altered by the spectral shape of sound stimulation, and how this interplay might manifest in behaviour. To this end, we recorded the EEG in N= 14 participants. Listeners either detected ("attended to") or ignored faint target sounds embedded within different noise stimuli that were generated to exhibit different 1/f slopes ($1/f^0$, $1/f^{-1}$, $1/f^{-2}$, $1/f^{-3}$) in their modulation spectra. We find attention to be a prerequisite for the tracking of 1/f stimulus statistics and discuss its relevance for the processing of signals and noise.

P.30 Transcranial 10-Hz stimulation but also eye closure modulate auditory attention

Malte Wöstmann[1], Lea-Maria Schmitt[1], Johannes Vosskuhl[2], Christoph S. Herrmann[2] & Jonas Obleser[1]
[1] Department of Psychology, University of Lübeck, Germany
[2]Experimental Psychology Lab, Center for Excellence "Hearing4all," European Medical School, University of Oldenburg, Germany

When humans focus attention to auditory events, neural alpha oscillations (~10 Hz) in the Magneto-/Electroencephalogram (M/EEG) increase in power. Here we test whether experimentally induced increases in alpha power modulate auditory attention. In two studies, healthy human participants attended to spoken target digits against distractors. Alpha power was increased exogenously through transcranial alternating current stimulation (tACS), or endogenously through eye closure. In study I (*n* = 20), participants were cued to attend to a stream of four spoken digits presented to one ear, while ignoring a distracting (same-talker) stream of digits presented to the other ear. Previous M/EEG studies have shown that such dichotic tasks increase alpha power in auditory and parietal cortex ipsilateral to the focus of attention. To manipulate this alpha lateralization, we applied continuous 10-Hz tACS to temporal and parietal scalp regions in the left hemisphere (1 milliamp; stimulation sites FC5 and TP7). To control for the effect of stimulation frequency, each participant also received sham and gamma-tACS (47.1 Hz). Compared to sham, left-hemisphere alpha-tACS enhanced the recall of target digits in 'attend-left' versus 'attend-right' trials, while the opposite was found for gamma-tACS. This suggests that an exogenous increase in lateralized alpha power relatively suppresses auditory spatial attention to the side opposite to stimulation. In study II (*n* = 22), we sought to invoke an endogenous increase of alpha power instead while presenting participants with two alternating (different-talker) streams of five spoken digits. On each trial, participants were instructed to attend to one stream and to ignore the other. In blocks where they closed their eyes (in a dark room; compared to keeping their eyes open), participants induced a baseline increase in parieto-occipital EEG alpha power. During a trial, baseline-corrected alpha power fluctuated rhythmically, with alpha peaks preceding onsets of attended digits by ~100 ms. This attentional modulation of alpha power strongly increased with closed compared to open eyes, demonstrating that eye closure boosts the neural difference in auditory attending versus ignoring. However, eye closure did not enhance participants' ability to afterwards tell attended from ignored digits, which contradicts the widely held belief that eye closure enhances the behavioral outcome of attentive listening. In sum, the observed impact of eye closure and alpha-tACS on neural alpha dynamics and behavioral corollaries suggest that alpha power is more than a mere epiphenomenon but neurally and behaviorally relevant to auditory attention.

Main railway station

Arriving at the Atlantic Hotel:

15 min

5 min, stop Sandstraße
Line 1, direction Hochschulstadtteil
Line 2, direction Sudetenstraße/Bornkamp
Line 6, direction Blankensee
Line 9, direction Grillenweg

ZOB

Atlantic Hotel
Schmiedestraße 9–15

Buddha Bowl
Schmiedestraße 26

Gang No. 56
Marlesgrube 58