BAS BASIC AUDITORY SCIENCE 2023

Programme

Imperial College London

September 21st – 22nd

Organising committee

Lorenzo Picinali

Andrei Kozlov

Aidan Hogg

Katarina Poole

Programme

Day 1: Thursday, 21 September

Each session will consist of 4 presentations (10 mins each, leaving 5 mins for questions)

09:00 – 09:45 Registration & Coffee

09:45 – 10:00 Welcome Session

10:00 – 11:00 Session 1: Hearing-Assistive Technologies and Modelling - Chair: Mark Fletcher

- An interactive evaluation of gaze-directed beamforming in noisy conversations - John Culling
- Towards real-world benefit with a cochlear implant speech coding strategy that leverages temporal masking effects - Lidea Shahidi
- Watching Hearing with a Neuro-Implant: Ultra-high Resolution Models of Neural Activity in the Human Inner Ear - Werner Hemmert
- Recovering speech intelligibility for cochlear implants with deep-learning strategies in noisy reverberant situations using one or more microphones - Clément Gaultier

11:00 – 11:30 Break + Posters

11:30 – 12:30 Session 2: Spatial Hearing and Immersive Audio Rendering - Chair: Lorenzo Picinali

- HRTF upsampling: A machine learning approach - Aidan Hogg
- Remote Acoustic Soundscape evaluation using 3rd order Ambisonics recordings - Balandino Di Donato
- Perceptual Evaluation of Low-Complexity Diffraction Models from a Single Edge and Efficient Diffraction Modelling using Neural Networks - Joshua Mannall
- Three dimensions of space in object-based audio: perceived source distance - Philip Jackson

12:30 – 14:00 Lunch + Posters

14:00 – 15:00 Session 4: Objective Measures of Hearing (EEG, fNIRS, fMRI etc.) - Chair: Chris Sumner

- Studying the neural basis of changes in the perceived loudness following the activation of a cochlear implant - Dorothée Arzounian
- Electrophysiological evidence for prediction errors during perception of degraded spoken sentences - Ediz Sohoglu
- Frequency mapping in human primary auditory cortex predicts frequency discrimination performance - Julien Besle
- Gaussian Processes for efficient audiogram estimation with Auditory Brainstem Responses - Steven Bell

15:00 – 16:00 Break + Posters

16:00 – 17:00 Session 3: The Physiology of Hearing - Chair: Kerry Walker

- Visual signals in ferret auditory cortex - Rebecca Norris
- Epoch dependent encoding of category, decision, and reward in localized regions of ferret frontal cortex - Jeffrey Boucher
- Genetic risk for schizophrenia and experience of hearing impairment both influence auditory brain function in a mouse model of 22q11.2 Deletion Syndrome - Chen Lu
- Environmental noise modulates sound encoding in adult rat auditory cortical neurons - Natsumi Homma

19:00 – 22:30 BAS Dinner

Day 2: Friday, 22 September

10:00 – 11:00 Posters

11:00 – 12:00 Session 5: Psychoacoustics and Cognitive Aspects of Normal and Impaired Hearing - Chair: Brian Moore

- Pupil Dilation and Microsaccades Provide Complementary Insights into the Dynamics of Arousal and Instantaneous Attention during Effortful Listening - Maria Chait
- Relative pitch representations and invariance to timbre - Malinda J. McPherson
- Effects of hearing loss on listener's ability to detect mistuning in music - Sara Miay Kim Madsen
- An Association Between Auditory Responsiveness of Children and Duration of Entertainment Screen Time in the First Two Years of Life - Josephine Marriage

12:00 – 13:30 Lunch + Posters

13:30 – 14:30 Session 6: Multi-Modal Hearing and Hearing-Assistive Technologies - Chair: Amir Hussain

- Developing a haptic hearing-assistive device to improve speech perception in hearing-impaired listeners - Samuel Perry
- Searching for individual differences in audiovisual integration of speech in noise - Chris Sumner
- How does vision benefit speech perception in noise? - Lida Alampounti
- End-to-End demonstration of an Audio-Visual Speech Enhancement Model in an Embedded Edge AI System - Islam Zakaria Nasr

14:30 – 15:00 Break + Posters

15:00 – 16:00 Session 7: Effects of Noise Exposure on Hearing and Therapeutic Training - Chair: Chris Plack

- Hearing Assessment of Music Conservatoire Students - Stephen Dance
- Blood prestin levels following music exposure that induces temporary threshold shifts - Eleftheria Iliadou
- Do auditory brainstem responses reflect cochlear synaptopathy - evidence from frequency-specific responses - Katrin Krumbholz
- Perception and Measurement of Audio Quality Attributes in Music by Hearing Aid Users: A Sensory Evaluation Study - Alinka Greasley

16:00 – 16:10 BAS Closing Session

16:10 – 16:30 Break

16:30 – 18:00 UK Acoustics Network Plus satellite event on the Future of Acoustics for Wellbeing and Health (all welcome)

Information

Oral presentations

These will be presented in G20 in the Royal School of Mines and should be 10 minutes long. Presentation slots are 15 minutes in total to allow for 5 minutes for questions.

Poster presentations

These will be presented in G01 in the Royal School of Mines and should be a maximum size of A0 and should be in portrait format.

Venue

The venue for the meeting is in Lecture Theatre G20 for the oral presentations and G01 for the poster presentations which are located on the ground floor of the Royal School of Mines (Royal School of Mines, South Kensington Campus, London, SW7 2AZ) opposite the Royal Albert Hall, with the closest tube station being South Kensington.

The easiest way to access the building is off Prince Consort Road and entering the building from there (see red arrow on the map on the next page). There will be signage then directing you to the two lecture theatres.

Wifi at the venue is provided via eduroam

Transport

The best way to get to the venue is by tube or bus and both accept contactless cards as payment. We highly recommend using the app 'Citymapper' to navigate London as it gives mostly accurate timings for bus and tube times.

Registration/Refreshments

These will be provided in G01 in the Royal School of Mines. Coffee and tea is included in the registration cost

Lunch

Lunch is not included in the registration cost but there are several places on campus and in South Kensington where it can be purchased. Some recommendations are:

- Fusion 54 in the Dyson Building
- Library café in the Central Library
- The Royal School of Mines Café in the Royal School of Mines
- h-bar on level 0 of the Sherfield building
- Verdi Italian Kitchen in the Royal Albert Hall
- Berlin on Exhibition Road

Dinner

Dinner is included in the registration cost and will be a fork buffet (vegetarian and vegan options available) in Queens Tower Rooms located in the Sherfield Building. The easiest way to get there is via Dangoor Plaza, noted with the purple arrow on the map on the next page. Otherwise, you can also access the Sherfield Building via the walkway and go down one floor to get to Queens Tower Rooms. Signage will be able to direct you.

Accommodation

The meeting will start at 9am and finish at 5pm, followed by a conference dinner, on the 21st of September and start at 10am and finish at 6pm on the 22nd of September. We therefore recommend booking accommodation for the night of the 20th and 21st of September if accommodation is needed. Unfortunately, accommodation cannot be provided through the university, therefore we have provided a list of recommendations below.

If guests quote 'Imperial College' a discount may be applied.

- The Baileys Hotel London – reservations.baileys@millenniumhotels.com – 0207 331 6196
- Chelsea Cloisters – reservations@chelseacloisters.co.uk – 0207 584 1002
- Fraser Suites Queens Gate – sales.london@frasershospitality.com – 0207 341 5599
- The Gore Hotel – reservations.thegore@starhotels.com – 0207 590 6704

- Grange Strathmore Hotel – Strathmore.reservations@gemhotels.com – 0207 584 0512
- Holiday Inn Express Heathrow T5 – reservations@hiexheathrowt5.com – 01753 684001
- Ibis Shepherds Bush – h7813-re@accor.com – 0207 348 2020
- The Mercure Paddington – Christina@lth-hotels.com – 0207 244 2400
- The Park International Hotel – reservations@parkinternationalhotel.com – 0207 244 4677
- 100 Queens Gate – reservations@100queensgate.com – 0207 400 3800
- Radisson Blu Edwardian Vanderbilt – resvand@radisson.com – 0207 761 9000
- The Rembrandt Hotel – reservations_rembrandt@sarova.com – 0207 589 8100

Finally, we'd like to thank the UK Acoustics Network Plus (UKAN+ – https://acoustics.ac.uk/) and EPSRC for funding and supporting this event.

If at any point you require assistance, please don't hesitate to ask a member of the organising committee.

BAS2023 Imperial Organising Committee:

Lorenzo Picinali,
Andrei Kozlov,
Aidan Hogg,
Katarina Poole

Session 1: Hearing-Assistive Technologies and Modelling

Chair: Mark Fletcher (University of Southampton)

Oral presentations

O1.1 An interactive evaluation of gaze-directed beamforming in noisy conversations.

John Culling, Niamh Powell, Bryn D. Davies, Emilie F. C. D'Olne & Patrick A. Naylor.

Cardiff University, UK.

Gaze-directed beamforming has potential for future hearing aids by selecting sound from the direction of the user's gaze. We evaluated this potential using a simulation of an 8-microphone array mounted on spectacles. In phase 1, participants watched 6 segments (165 seconds each) of a zoom call between two parties located at ±15˚ with interfering talkers at 0˚, ±60˚, ±135˚ and 180˚. One condition simulated binaural hearing over headphones using head-related transfer functions. The other condition used an eye-tracker to select filters for each source and for each video frame from a look-up table. The table was based on the predicted directionality of the microphone array in diffuse noise for a minimum-variance distortionless response beamformer. The participant's gaze and the headphone output were recorded. Participants completed a short questionnaire about the conversation after listening to each segment. In phase 2, the intelligibility of each sentence was measured formally. Each recording was presented to a new, yoked participant, sentence-by-sentence for transcription. An individually tailored video cut back and forth between the two talkers in accordance with the eye-tracking record of the corresponding phase-1 participant in order to provide identical lip-reading cues. In phase 1, the scores on the questionnaire were approximately doubled by use of the beamformer and, in phase 2, the number of words correctly transcribed also doubled.

O1.2 Towards real-world benefit with a cochlear implant speech coding strategy that leverages temporal masking effects

Lidea Shahidi, Robert P. Carlyon, Deborah A. Vickers & Tobias Goehring

University of Cambridge, Cambridge, UK

Speech perception remains challenging for cochlear-implant (CI) recipients in listening conditions containing background noise. The temporal integrator processing strategy (TIPS) can significantly improve the perception of CIS-processed speech in the presence of speech-shaped noise (Lamping et al. 2020). TIPS uses a model of temporal masking to identify and remove stimulation pulses that are unlikely to be perceived. Here, the potential for TIPS to benefit real-world speech perception is investigated through adaptation of the algorithm for real-time processing (a key constraint of CI devices), and computational analyses and speech perception testing in everyday listening conditions. The computational analyses indicate that the pulse-removal behavior of TIPS is consistent across everyday listening conditions but varies with the listener's stimulation map parameters. In all conditions, TIPS could considerably reduce the power required for stimulation, addressing a major limitation of current devices. A low-latency semi-causal version of TIPS was developed and validated

in simulation, demonstrating the feasibility of applying TIPS in clinically-used devices. We will discuss results from the ongoing listening test, revealing whether additional speech perception benefits noted in the computational analysis occur in everyday listening scenarios.

Lamping W, Goehring T, Marozeau J, & Carlyon, RP. Hearing research. 2020; 10.1016/j.heares.2020.107969.

O1.3 Watching Hearing with a Neuro-Implant: Ultra-high Resolution Models of Neural Activity in the Human Inner Ear

Werner Hemmert[1], Albert Croner[1], Rudolf Glückert[2], Anneliese Schrott-Fischer[2] & Siwei Bai[1]

[1]Technical University of Munich, Germany; [2]University of Innsbruck, Austria

Cochlear implants (CIs) are the most successful neuroprostheses: they restore hearing in deaf people to a surprisingly high degree. We have developed models of the inner ear from high resolution µCT scans (down to 6 µm voxel size) from cadaveric human temporal bonesand reconstructed the path of up to 400 individual fibers of the auditory nerve. We have virtually inserted an electrode array and calculated the electrical current spread with a finite element model (22 Mio. elements). With a biophysically motivated multi-compartment model, which included the most important voltage-activated ion channels, we were able to derive excitation patterns of the auditory nerve fibers along the cochlea. We have also modeled neuronal adaptation in the human auditory nerve to electrical stimulation with a cochlear implant with two mechanisms: slow inactivation of sodium channels and M-type potassium currents.

In summary, our model predicts both neuronal excitation patterns and neuronal spike rate adaptation in the human auditory nerve for electrical stimulation. This is a major breakthrough to our understanding of the complex, superimposing nonlinear dynamics of neurons in general and especially relevant for all types of electrical and optogenetic stimulation. With these predictions it is possible to devise the next generation of CIs and coding strategies. In addition, our spectacular visualizations provide a detailed insight into the function of the most delicate sensory organ.

O1.4 Recovering speech intelligibility for cochlear implants with deep-learning strategies in noisy reverberant situations using one or more microphones

Clément Gaultier, Tobias Goehring

University of Cambridge, Cambridge, UK

Noise and reverberation can impair speech perception for cochlear implant (CI) recipients. With improvements in end-to-end deep learning techniques and multi-microphone approaches, powerful speech enhancement (SE) strategies can now be developed to jointly mitigate noise and reverberation in challenging listening situations. Three algorithms were trained to jointly remove noise and reverberation on simulated sound scenes using either single- or multi-microphone recordings from behind-the-ear devices. Performance was assessed using objective measures and a listening test with 12 CI listeners and 15 participants with typical hearing using CI simulations. All approaches showed improved objective scores over the noisy reverberant mixtures. Metrics indicated better signal to distortion ratios and predicted intelligibility for the multi-microphone conditions, especially in the noisiest situations. Experimental results showed significant improvements in Speech Reception Thresholds for the multi-microphone approaches over both the noisy reverberant and single-microphone cases. Multi-microphone SE algorithms based on deep learning showed strong potential to improve speech intelligibility in realistic situations with babble noise and reverberation for CI listeners. This was the case even when restricting the processing to unilateral microphones. Further work should investigate the effect of such strategies on preserving auditory awareness of the environment whilst enhancing the speech.

Poster presentations

P1.1 Simultaneous multiple beam-steering in hearing aids for speech enhancement in multi-talker communication scenarios

Hendrik Kayser[1], Daniela Rothenaicher[1,2], Giso Grimm[1] & Volker Hohmann[1]

[1]*University of Oldenburg, Oldenburg, Germany; [2] now with Vibrosonic GmbH, Mannheim*

Everyday communication includes scenarios with multiple talkers at different locations and background noise, which remain challenging for hearing-impaired people even when using hearing aids. To improve speech intelligibility in such scenes, we propose directional processing with multiple steerable beamformers based on the direction of arrival of each talker of interest. In addition to the noise reduction, simultaneous beams may reduce the detrimental effect of too-slow adaptations after fast turn-taking of target talkers, which is potentially limiting the benefit of single-beam approaches. In this study, virtual acoustics, real-time hearing device processing, and realistic head motion data were combined to assess multiple-beam approaches for signal enhancement in hearing aids. The approaches under test included standard hearing aid algorithms as well as optimal directional filtering in the virtual acoustic rendering system (oracle knowledge). Algorithms were tested in different acoustic scenarios, varying from simple conditions with two target talkers to scenes with multiple target and interfering talkers in a noisy reverberant environment. Simulated parametric head rotations as well as measured real movements of subjects were used. In our simulations we found that a combination of multiple instances of a binaural minimum-variance distortionless response beamformer, forming multiple simultaneous beams, resulted in a benefit over prevalent single-beam hearing aid algorithms.

P1.2 A “hilltop” approach for tackling channel interactions in Cochlear Implant users

Francois Guerit[1], Iwan Roberts[2], Chloe Swords[2] & Robert P. Carlyon[1]

[1]*Cambridge Hearing Group, MRC Cognition and Brain Sciences Unit, University of Cambridge;* [2]*Cambridge Hearing Group, Department of Clinical Neurosciences, University of Cambridge*

Cochlear Implant (CI) users struggle when listening to speech under noisy conditions. This has long been attributed to channel interactions. Attempts at reducing them with sharp current focusing (e.g. tripolar) mostly failed, reflecting that the outputs of adjacent channels are highly correlated, and that some neural spread of excitation is necessary for sufficient loudness. We pilot a “Hilltop” approach, where the voltage is shaped by a Gaussian function (with σ = 2, 3, and 4 elec.) to be broad near a given electrode, but reduced by several orders of magnitude more than 3-4 electrodes away.

In each participant, using a constrained least-squares optimization, we obtain the current amplitude and polarity to be injected at each electrode to produce the desired Hilltop voltage patterns. Participants balanced the loudness of Hilltop stimuli to a monopolar (MP) stimulus, for three different centre electrodes. We also measured charge summation as a function of inter-channel spacing, and are currently investigating channel pitch ranking.

At equal loudness, Hilltop stimulation produced large voltage reductions compared to MP stimulation more than 3 elec. away from the centre electrode. Charge summation results indicate a possible polarity reversal along the electrode array, which we are investigating with a 3D finite-element model of the cochlea. We expect the Hilltop approach to provide insights into the optimisation of electrical stimulation for reducing channel interactions.

P1.3 Assistive Listening Systems – the need for Standardisation and an Installation Code of Practice

Peter Mapp

Peter Mapp Associates Colchester UK

In the past few years, the technologies available to create wide area and public Assistive Listening Systems (ALS) have developed and increased potential choice. Whilst this has enabled a greater range of buildings, rooms and venues to be provided with ALS not all of these systems have proved to be successful when actually installed in real world environments. Whereas traditional hearing loop systems have, in general served their intended audiences well (particularly since the introduction of BS 7594 Code of practice for audio-frequency Induction loop systems (AFILS) in 1993) hearing loops have limitations and are always not suitable every purpose and venue. With the emergence of new technologies and potential release of Auracast devices and systems in the relatively near future, it is clear that some form of standardisation of an Assistive Listening System's acoustic and audio performance, together with a code of practice for installations is required. Furthermore, the responsibility of implementing and maintaining such systems is tending to migrate from audio-based engineers to IT based personnel, who often have little or no audio systems experience and training.

The paper sets out the basic performance requirements for wide area and public based ALS which potentially may be used to act as both industry and international codes if practice and performance standardisation.

P1.4 The benefits of bimodal hearing in children and adolescents: A systematic review and narrative synthesis

Yuhan Wong, Iain A Bruce, Josef Schlittenlacher, Karyn L Galvin & Karolina Kluk-de Kort.

Manchester Centre for Audiology and Deafness (Manchester, UK)

Background and aims: In the last two decades, cochlear implant (CI) candidacy criteria around the world have been expanded to include individuals with residual hearing. Unilateral CI recipients with residual hearing in the contralateral ear have been to benefit from a hearing aid in that ear (bimodal hearing). However, evidence of the real-life benefits of this combination of electric and acoustic hearing in the paediatric population varies. Moreover, there is a lack of evidence-based guidelines on whether a child with bimodal hearing should be considered for a second CI. In the UK, the latest National Institute for Health and Care Excellence (NICE) guidance for cochlear implantation in children (TAA566, 2019) has lowered the audiological criteria (pure tone audiometric threshold ≥80dBHL at 2 or more frequencies (500Hz, 1000Hz, 2000Hz, 3000Hz and 4000Hz) and who do not receive adequate benefit from acoustic hearing aids) but additionally stipulates that hearing loss must be bilateral. As such, children and adolescents with asymmetrical hearing losses where one ear is outside of audiological criteria may not have access to a CI. We aim to systematically review the literature on real-life outcomes of children and adolescents with bimodal hearing and how they compare with other hearing configurations (unilateral and bilateral CI).

Methods: A protocol for this review was registered on PROSPERO. Four databases were used for the literature search – MEDLINE, EMBASE, CINAHL and Cochrane. A search strategy was designed and inclusion/exclusion criteria were developed using the population, intervention, comparison, outcome and study design (PICOS) framework. Searches were carried out using keywords and subject headings around the population of interest (children and young people/adolescents) and the intervention (bimodal hearing). The quality of the evidence was assessed using the GRADE criteria and risk of bias with the ROBINS-I tools.

Results: Database searches yielded 433 results and 219 unique records after removing duplicates. After title and abstract screening, 63 studies were included for full-text screening, and 18 were found to meet the criteria. The outcome domains reported were speech perception in noise, speech perception in quiet, language development, pitch discrimination or accuracy, music perception, spatial hearing, and noise tolerance. While bimodal hearing outperformed unilateral CI in all studies comparing the two conditions, comparisons between bimodal hearing and bilateral CI yielded varied results.

Conclusion: This systematic review will aid multidisciplinary CI teams in their decision-making for individuals with borderline CI candidacy by extending current knowledge about bimodal hearing. Findings from this systematic review, along with an earlier review on the benefits of hearing preservation cochlear implantation, will be used in the design of a study on the benefits of combining electrical and acoustic hearing in paediatric cochlear implant patients.

P1.5 Effect of frequency-to-place mismatch and frequency warp on speech and music sound quality in acoustic cochlear implant simulation

Louis Villejoubert, Lorenzo Picinali, Kathleen Faulkner & Debi Vickers

University of Cambridge

In more recent years a greater emphasis has been placed on improving the sound quality for individuals using cochlear implants (CIs). Indeed, many post-lingually deafened adults with CIs have speech recognition performance close to normal hearing (NH) listeners in quiet environments, but yet complain about poor sound quality, especially for music. The contribution of CI technology to the degradation of music quality has been well documented, but little is known about the relative importance of different aspects relating to the technology and how they interact with the personal preferences and hearing characteristics of the individual. One individual factor that is suspected to contribute largely to the sound quality degradation is the frequency-to-place mismatch (FTPM) induced by the incomplete insertion of the electrode array inside the cochlea. FTPM is defined as a deviation between the characteristic frequency and the allocation frequency assigned during CI fitting.

Whereas much of the speech perception improvement occurs during the first months after CI activation, recent studies have shown that a significant FTPM could negatively affect this adaptation process. Alongside the improvements in speech perception over time from device activation patients often also change their reports of sound quality, from “high pitched”, “screechy”, “dalek-like” to more “normal” or “natural” after extensive listening experience. However, the specific contribution of FTPM on sound quality isn’t clear.

This research investigates the link between sound quality and FTPM in acute listening experiments with CI simulations with NH listeners situation and how pre-modifications of speech and music material can be used to explore and understand the impact of FTPM.

Testing used the MUlti Stimulus test with Hidden Reference and Anchor (MUSHRA) paradigm for listening to different acoustic simulations of monaural CIs for different sound materials (speech, vocal music with instruments, and instrumental only). The first experiment explored the link between sound quality and FTPM, where participants had to assess the sound quality of different FTPM profiles relative to a non-shifted reference. The second experiment was a frequency shifting experiment, where NH listeners compared and rated their listening experiences for the speech and music stimuli previously shifted.

Preliminary results from the first experiment indicate that FTPM degrades sound quality, with the level of degradation depending on the frequency content of the sound material and the frequency band in which the FTPM is simulated. However, for now no specific trend has been observed in the second experiment between the FTPM simulated and the degree of frequency that results in the best sound quality judgements.

This work will be continued with CI listeners with the intention of quantifying the degree of FTPM and the importance with the intention of informing re-mapping.

P1.6 Exploring objective measures of auditory temporal resolution

Esma Akis, Steven L Bell & David M Simpson

Institute Of Sound & Vibration Research, University of Southampton

Auditory temporal resolution is an important aspect of suprathreshold hearing, especially for speech comprehension and it is typically defined as the ability to detect temporal changes in a signal. Several methods have been proposed for the objective measurement of temporal resolution, although these approaches have not generally been compared in the same subjects. The present study aims to develop a reliable objective method to measure temporal resolution thresholds based on Auditory Brainstem Responses (ABR) combined with sensitive statistical signal detection methods.

This study included 23 adults (20-35 years) with normal hearing. To determine the objective temporal resolution thresholds, two approaches have been explored: 1. ABRs to temporal notched noise with clicks inserted in the middle of a variable duration gap 2. Two clicks with a variable gap between them. Bootstrap, Fsp, Hotelling's T2, and peak-to-peak amplitude estimation methods were used to detect the presence of ABR responses.

Most of the participants showed an ABR response for 4-ms and above gap durations. At a group level, the amplitude of the responses and the number of ABRs detected decreased as gap durations approached the threshold levels.

Both experimental paradigms with objective detection methods look promising to estimate temporal resolution at a group level, with similar performance as gap durations increase. Individual objective temporal resolution threshold determination appears more challenging.

P1.7 That looks noisy: Using visual cues to improve clinical assessment of self-reported hearing

Emanuele Perugia, Samuel Couth & Gabrielle H. Saunders

Manchester Centre for Audiology and Deafness (Manchester, UK)

Clinical audiological assessment and follow-up appointments often have patients complete a self-report questionnaire to evaluate hearing difficulties and hearing aid outcomes and benefits. There are many questionnaires available and all use a similar format, that of having the respondent rate their listening difficulty based on a written-description of a listening situation (e.g. understanding speech in a pub, or meeting).

Based on the principle that "a picture is worth a thousand words", we propose to create an image-based questionnaire that uses photographs instead of written descriptions of listening situations to assess self-reported hearing. Using photographs will increase accessibility for individuals with low literacy or who speak a different language.

The development of the questionnaire involved selecting photographs that were relevant, inclusive, and based on the Common Sound Scenarios framework (Wolters et al., 2016). This was achieved in a 3-part validation. Initially, a focus group was held with general public who provided input on the photographs. Then, a bidirectional validation study was conducted online. In Part 1, participants selected which of 6 photographs (if any) they considered best reflected the written description of a listening scenario. In part 2, using the best match photographs from Part 1, participants were asked to describe the listening or hearing scenarios portrayed by the photographs. The results of the bidirectional study will be presented.

P1.8 Gated-recurrent neural networks as encoding models of auditory cortical computation

Lorenzo Mazzaschi, Nicol Harper, Ben Willmore & Andrew King

University of Oxford, Oxford, United Kingdom

Computational models of the sensory brain are an invaluable tool to interpret complex neural responses to dynamic natural stimuli. Yet, much of the activity of the brain in response to such stimuli, and consequently the processing along the auditory pathway that gives rise to it, remains to be explained. In this project, we explore ways to use and improve computational modelling for understanding computations in the auditory brain. Current state-of-the-art encoding model architectures, stemming from advances in machine learning, leverage deep feedforward neural network structures to predict neural responses to natural sounds. However, such models can only account for approximately 70% of the response variance in neurons recorded from A1 of anesthetised ferrets, and 30% in awake ferrets. We set out to improve on these figures by introducing recurrent (cyclic) connections in these network models. As recurrency can be considered a in silico equivalent of in vivo memory and feedback, we hypothesized that the response prediction in awake neurons would particularly benefit from this addition. We indeed found that recurrency yielded an improvement in comparison with feedforward networks, and that more complex forms of recurrency (in which so-called gates selectively determine what information is to be kept in working memory) produced a further improvement, suggesting a key role for memory in perception, and a gate-like operation for a portion of cells in the auditory cortex.

P1.9 An inclusive digit-in-noise test with expanded applicability for auditory research and clinical applications

Thu Ngan Dang[1], Shangqiguo Wang[2], Clément Gaultier[1] & Tobias Goehring[1]

[1]MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, United Kingdom; [2]University of Hong Kong

People with hearing loss struggle most with hearing speech in background noise. Digit-in-noise (DIN) tests have been widely used as an index for hearing loss and to probe auditory function and speech perception in noise. Here we developed an inclusive DIN test (iDIN) to be used in auditory research and audiology, including people who use cochlear implants (CI). In this iDIN test version, recordings of male and female native British English speakers were obtained in a quiet environment and in different levels of background noise to elicit the Lombard effect. Native English speakers then rated the recordings to select the most natural stimuli. Responses from 26 typically-hearing young adults, equally divided into three groups (10 non-native, 8 native British and 8 native non-British English speakers), were obtained to extract speech reception threshold (SRT) levels across digits. We found consistent patterns of SRT and no significant differences between the 3 groups. Male speech in both conditions and female speech in the quiet condition generated similar SRT trends across digits, whereas female speech in the noisy condition produced a different SRT pattern with an outlier at digit 5. These results suggest that the iDIN test can be used for a wide population, including native and non-native English speakers, with further validation ongoing for CI users. The iDIN is aimed to provide a quick and accessible tool for clinical and research settings.

P1.10 Pupil dilation response as an objective index of 1/f behaviour of coloured noises

Mercede Erfanian, Maria Chait & Jian Kang

UCL (London, UK)

Coloured noises are sound signals whose power spectra result from stochastic processes. They have wide-ranging applications for enhancing memory consolidation, sleep quality, and attentional mechanisms, among others. The main feature of coloured noises is the dissimilarities in their power spectral density (PSD), exhibiting distinct 1/f structures which may serve as a potential indicator of varying levels of arousal. However, the physiological underpinnings of coloured noises remains unclear and requires further investigation. To understand the underlying physiological alterations, we evaluated pupil dilation response (PDR) as a measure of tonic arousal. To this end, we generated auditory stimuli consisting of white, pink, and brown noises. In an experiment, participants listened to several coloured noises while their pupil size dynamics were recorded. The hypothesis posited that white noise would elicit a greater pupil size, reflecting its heightened level of arousal, followed by pink and brown noises, respectively. The results showed that the pupil size was not modulated by the coloured noises. These findings suggest that despite wide ranging hypotheses regarding different arousal nature of colour noises, they do not evoke differing sustained pupil diameter responses.

P1.11 The 2nd Clarity Prediction Challenge For Hearing Aid Speech Intelligibility

Trevor Cox, Will Bailey, Simone Graetzer, Jon Barker, Michael A. Akeroyd, Graham Naylor & John F. Culling.

University of Salford

This paper reports on the design and outcomes of the 2nd Clarity Prediction Challenge (CPC2). To develop better hearing devices for individuals, we need automatic ways to predict personalised speech intelligibility, and therefore need models that can take as parameters such as hearing aid and listener characteristics. The Clarity project's methodology is to use open signal processing challenges to harness multiple prediction approaches from different research groups. In CPC2, the hearing aid audio came from processing speech in the presence of typical domestic noise in a living room. Each signal contained a short sentence and listeners were asked to repeat the words. Challenge entrants had to build a system to predict how many of the words were recognised correctly by the listeners. Prediction models could use the clean speech and take an intrusive approach; alternatively they could be non-intrusive and just use the hearing aid signals. We will review the algorithms that entrants used, and compare performance across approaches.

P1.12 Remote from switch on – a pilot case study for innovative service delivery in adult cochlear implants

Dakota Bysouth-Young[1], Mohammad Shadid[2] & Andrew Chang[3]

[1]MRC Brain and Cognition Sciences (Cambridge, UK); [2]Hearing Implants Australia; [3]Mater Health Services Brisbane

Improvements in technology and the growing acceptance of tele-services have implications for innovating service delivery models. We examine the effectiveness of providing post-operative audiological cochlear implant services remotely. We present a detailed case study on delivering acute CI services remotely, starting from the switch-on stage. We describe the framework for coordinated service delivery involving multiple independent stakeholders and discuss how audiology services were adapted using telehealth solutions to ensure positive outcomes for recipients. Additionally, we explore the future implications for outcomes and effective service delivery models that are device manufacturer-agnostic. Post-operative speech and functional outcomes showed significant improvements compared to pre-operative measures, with results within the expected outcomes for cochlear implant recipients. The cost of service delivery was comparable to standard procedures but offered recipients greater flexibility. We discuss the implications of these findings in relation to emerging technologies and changes in reimbursement practices in Australia. This study paves the way for hearing healthcare providers to explore innovative approaches in meeting the demand for CI intervention among individuals with significant hearing loss. By innovating service delivery models, we can potentially eliminate barriers to access care. Further research is necessary to validate the efficacy of this approach.

P1.13 Feasibility study for the Nottingham Hearing and tinnitus BioResource

Ruth Spriggs[1,2], Paul Bateman[1], Raul Sanchez Lopez[1,2], Derek J. Hoare[1,2] & Ian M. Wiggins[1,2]

[1]NIHR Nottingham BRC, Nottingham, United Kingdom; [2]University of Nottingham, United Kingdom

Nottingham BRC has long term plans to establish a BioResource, a biomedical database used for research, to help define the next generation of audiological healthcare. This project is a feasibility study for the BioResource; a small sample of participants will be recruited to ensure that the processes put in place are all working optimally. Participants with and without hearing problems will undergo a series of audiological tests, give a hair sample (if appropriate) and complete hearing and general health questionnaires. Inclusivity will be important, as translational research should involve a representative sample of the population living with the condition, in this case hearing loss or tinnitus. The outcomes of this study will help the Hearing and Tinnitus BioResource team understand the challenges and innovations required to deliver a large scale programme. The aim of the BioResource is to follow participants as they age, by periodically repeating the tests and questionnaires, to allow researchers to link together different aspects of data, identify long term patterns, and ultimately improve treatment and quality of life for patients.

P1.14 Development of AUDITO: a web-based listening test system and database for inclusive hearing research

Tobias Goehring, Lidea Shahidi, Alexis Deighton MacIntyre, Thu Ngan Dang, Clement Gaultier & Bob Carlyon

University of Cambridge

Recently, web-based applications in audiology and hearing research have seen strong demand and development, driven by technological progress and limited in-person contact. We developed a web-based system to administer various listening tests either in the lab or remotely for increased inclusion in cochlear implant (CI) research and collaborations between researchers at different institutions. It capitalises on connectivity features using wireless streaming for cochlear implant technology, which is becoming available for most CI speech processors. A novel automatic connection check was implemented to test whether streaming is active during the listening test. To complement the listening test system, an acoustic database is under development with more realistic and inclusive auditory stimuli. Lessons learned from the development and first results from validation and research studies will be presented. Comparisons of interest include the presentation of auditory stimuli via direct input vs Bluetooth streaming, as well as in-lab vs at-home data collection. A web-based research system to administer auditory perception tasks via streaming technology in CIs or via headphones provides an opportunity to increase inclusiveness (no travel required), to save time and financial costs and to facilitate data collection for obtaining larger or longitudinal samples. This system is a promising way to improve ecological validity, inclusion and real-world translation of CI research.

P1.15 Investigating the effectiveness of Apple AirPods Pro as low-cost personal sound amplification devices

Stefan Bleek & Mariyam

University of Southampton

A subset of individuals with normal audiograms report hearing difficulties, particularly in challenging listening environments. The current management approach lacks a standardized protocol, resulting in varied clinical practices. Over-the-counter (OTC) amplification devices, including PSAs like AirPods Pro, have the potential to be a cost-effective alternative for individuals with normal and mild HL who report of hearing difficulties.

Preliminary findings from Joaquin et al. (2023) have shown promising results in improving speech perception and reducing mental effort for individuals with normal audiograms using Apple AirPods Pro.

This ongoing study aims to evaluate the effectiveness of Apple AirPods Pro (2nd generation) as low-cost personal sound amplification (PSA) devices for individuals with normal and mild HL with hearing difficulties. Participants with normal audiograms and mild HL will be recruited. Using the AirPods headphone accommodation feature, we will tailor the amplification according to each participant's individual audiogram. The study will use the Oldenburg matrix test to measure speech perception in noise. A questionnaire survey will be carried out to evaluate clinicians' subjective experiences in dealing with such individual cases. The study also addresses the affordability aspect of assistive listening devices, as traditional HAs are often not covered by health insurance and can be expensive.

The results of this study will provide valuable insights into the performance and efficacy of AirPods Pro as low-cost PSAs for individuals with normal or mild HL. It will contribute to the growing body of research exploring the use of consumer electronic devices for hearing assistance. The findings may have implications for improving accessibility to hearing assistance devices and addressing the unmet needs of individuals with normal and mild HL who experience difficulties in challenging listening situations.

P1.16 Advancements in Speech Enhancement Using Deep Neural Networks: An Analysis of Supervised, Unsupervised, and Semi-supervised Learning Approaches

Jianqiao Cui & Stefan Bleeck

Southampton University

Deep neural networks have significantly transformed the field of speech enhancement by leveraging the power of deep learning. While supervised learning relies on paired data for model training, challenges arise from issues such as limited diversity in the training dataset and the potential bias introduced by over-represented groups in public datasets. Furthermore, the performance of supervised models may be hindered when faced with unseen target speakers or environments due to the scarcity of paired data across all possible conditions.

To circumvent the formidable challenge of collecting real-world data, simulated data is often used, but this introduce a mismatch between the training and testing conditions, leading to inconsistent performance. Unsupervised learning methods offer an alternative by leveraging unpaired data for training, thereby avoiding the limitations associated with paired datasets.

Recent advances have explored semi-supervised and generative adversarial network (GAN) approaches. We propose here a novel speech enhancement approach that leverages a large volume of unpaired noisy and clean speech data for self-supervised pre-training. Subsequently, a limited amount of simulated paired data is used for fine-tuning the pre-trained model during the training phase. Experimental results demonstrate that the proposed method achieves superior performance compared to other unsupervised learning methods, highlighting its efficacy in speech enhancement tasks.

P1.17 Crosstalk cancellation for bilateral bone-anchored hearing aids.

Ryan Barnsley, John F. Culling & Robert J. W. Mcleod

Cardiff, U.K.

Bilateral bone-anchored hearing aids (BAHAs) suffer from crosstalk. This problem could be alleviated using cross-talk cancellation. Previous research used a tone from one bone transducer (BT) to cancel the crosstalk from a contralateral BT. Measurements at a range of frequencies were used to cancel crosstalk from a broadband noise and improve detection thresholds for tones presented on the cancelled side. The current research extended this procedure by applying cancellation bilaterally and simultaneously for complete crosstalk cancellation. At the start of every session, the phase and level of tones were adjusted until they cancelled the crosstalk from a contralateral BT, resulting in a strongly lateralised percept. Bilateral crosstalk-cancellation filters were then designed using this data. Masked thresholds for a tone at one BT and noise at the other were measured using the crosstalk-cancellation filters. Participants showed benefits of crosstalk cancellation in the range of 7-17 dB. Most participants were able to do the experiment after training. Currently we are in the process of recruiting patients with osseo-integrated bilateral BAHAs to test using the same procedure. This research has two goals: first to determine whether the results seen in normal hearing participants are replicable in bilaterally implanted patients using their abutments, and second to assess how stable the filter calibration is when using an abutment.

Session 2: Spatial Hearing and Immersive audio Rendering

Chair: Lorenzo Picinali (Imperial College London)

Oral presentations

O2.1 HRTF upsampling: A machine learning approach

Aidan Hogg, Mads Jenkins, He Liu & Lorenzo Picinali

Imperial College London

An individualised head-related transfer function (HRTF) is essential for creating realistic virtual reality (VR) and augmented reality (AR) environments, which could, in turn, be used for hearing research purposes. However, acoustically measuring high-quality HRTFs requires expensive equipment and an acoustic lab setting. To overcome these limitations and to make this measurement more efficient HRTF upsampling has been exploited in the past, where a high-resolution HRTF is created from a low-resolution one (i.e. one which could easily be measured at home). This work explores how machine learning can be used for this upsampling task. The advantage of machine learning approaches over traditional upsampling is that they are able to recreate the missing information in the sparse measurements using the knowledge learnt from a training set containing many high-resolution HRTFs. Results are given for an approach that transforms the HRTF data for convenient use with a convolutional super-resolution generative adversarial network (SRGAN).

O2.2 Remote Acoustic Soundscape evaluation using 3rd order Ambisonics recordings.

Balandino Di Donato & Iain McGregor

Edinburgh Napier University

The Acoustics - Soundscape ISO standard (ISO 12913-1:2014) has been used in previous work to assess individuals' soundscapes during controlled, limited soundwalks. There are several easily identifiable challenges to undertaking this type of approach, including logistics, and cost, along with potential participant generated noise. These issues, together with the environmental impact are greatly amplified if we want to realise a large-scale study across multiple locations. In this, pilot, we consider the potential of the Acoustics - Soundscape ISO guidelines for assessing recorded auditory environments remotely. We utilised a series of custom 3rd order Ambisonics recordings to run an evaluation of these auditory environments adapting the Acoustics - Soundscape ISO standard for a remote listening evaluation study. Third order ambisonics have the benefit of increased spatial accuracy compared to first order due to the total number of channels increasing from 4 to 16. In this presentation, we will present our approach and initial results from the pilot study, along with a discussion about how the method could be utilised within a wide variety of pertinent contexts, such as in augmented reality, virtual tourism, or ecoacoustics research.

O2.3 Perceptual Evaluation of Low-Complexity Diffraction Models from a Single Edge and Efficient Diffraction Modelling using Neural Networks

Joshua Mannall[1], Lauri Savioja[2], Orchisama Das[1], Paul Calamia[3], Russell Mason[1] & Enzo De Sena[1]

[1]University of Surrey, UK; [2]Aaloto University, Finland; [3]Reality Labs Research at Meta

Plausible real-time acoustic simulation is an ongoing research field. Many real-time applications use geometric acoustic models which model sound propagation as rays. This fails to model wave-like properties of sound such as diffraction. Diffraction can be modelled by introducing diffracted rays. Common models include the physically accurate Biot-Tolstoy-Medwin (BTM) model and the Uniform Theory of Diffraction (UTD). Recently, efficient models have been proposed to approximate BTM and UTD. We carry out a perceptual experiment to evaluate the perceived naturalness of BTM, UTD, the Volumetric Diffraction and Transmission (VDaT) model, an infinite impulse response (IIR) filter model (IIR-LO) and a proposed higher-order IIR filter model (IIR-HI). Stationary and moving receivers were considered for a single wedge in free field. The results suggest that IIR-HI is perceptually similar to BTM. VDaT and IIR-LO were found to be less natural in some cases. In dynamic scenes, VDaT was found to be more natural than the other models. The experiment was limited in scope by the simplicity of the scenes. However, the results suggest the approximate models are perceptually similar to the physically motivated models. We then use advances in machine learning to propose an IIR model trained on data generated using BTM. We show that our model approximates BTM with a mean absolute error of 1.0 dB for 1st-order diffraction and a lower computational cost than the current state of the art using UTD.

O2.4 Three dimensions of space in object-based audio: perceived source distance

Philip Jackson, Craig Cieciura & Afshan Farheen

University of Surrey

An impression of space and the distance of a sound source are arguably more salient auditory-scene descriptors than its azimuth of elevation angles. As the radial dimension in object-based audio, we studied the perceived distance of a phantom source reproduced over loudspeakers. Spatial audio reproduction of the source was achieved via the Reverberant Spatial Audio Object (RSAO), which parameterises an acoustical room impulse response (RIR) from source to listener, representing the direct path, early reflections and late reverberation as a set of diffuse and point source objects. We tested rooms from the SurRoom dataset, used previously to evaluate plausilbility of the RSAO's room impression over binaural reproduction. We focussed on RIRs recorded across a series of distances, from 1m to 3m, and used their RSAO parameters to create stimuli with corresponding phantom sources via the EBU ADM Renderer (EAR; European Broadcast Union, Audio Definition Model). Participatory listening tests in a sound treated room over a 22.1 setup in the Surrey Sound Sphere gathered ratings of apparent source distance using a MUSHRA-style user interface. Experimental results show the RSAO's strong effect on participants' perception of source distance, through a clear trend in each room, confirmed by statistical significance tests. Future manipulation of the RSAO parameters to interpolate and extrapolate distance measurements facilitates creative uses with source and listener motion.

Poster presentations

P2.1 Neural Correlates of Change Detection in Complex Spatialized Auditory Scenes

Drew Cappotto[1], Maria Chait[1], Vincent Martin[2], Lorenzo Picinali[2], Katarina C. Poole[2] & Martha Shiell[3]

[1]University College London; [2]Imperial College London; [3]Eriksholm Research Centre

The auditory system is constantly engaged in assessing ongoing auditory scenes, providing crucial early detection of emerging sound sources. This study aims to investigate neural correlates of non-target spatial hearing, through the development of an ecologically-relevant measure of spatial auditory change detection. Human participants are presented with spatialized auditory scenes using discrete loudspeaker positions within a bespoke dome-shaped array. Scenes are composed of sounds ('sources') that maintain naturalistic features without conveying any specific semantic meaning to the listener. Each auditory scene consists of multiple pseudo-randomly spatialized 'sources', with an additional source added midway from the left, right, front, above, or behind the participant.

Analysis of participants' ($n = 11$) behavioral responses reveals proficiency in detecting source appearance (mean across locations and sources: 90.18% accuracy, 865 ms reaction time), with significant decreases in performance as the number of simultaneous masker sources increases from three to seven sources ($\beta = -3.364$, $p < 0.001$). In further investigations, we will analyze EEG data recorded in naive listeners, to identify event-related potentials (ERPs) that may be modulated by the position of added chimeras which may be suggestive of a neural code for change detection location in spatial auditory scenes. Furthermore we will employ this paradigm on a pared-down speaker array with phantom image locations.

P2.2 Simulated Hearing Loss Does Not Impair People's Ability to Echolocate

Denise Foresteire & Lore Thaler

Durham University (Durham, UK)

The current study addressed how echolocation in people might be affected by simulated hearing loss. Healthy-hearing, sighted participants listened to binaural recordings of echolocation sounds and performed a range of behavioural tasks: Echo Detection, Echo Localization and Navigation of a virtual echo-acoustic space. We found a significant impairment in performance in both echo-detection and echo-localisation tasks when we simulated hearing loss above 2 kHz, but not below 2 kHz, both bilaterally and unilaterally. Yet, most of the energy in our echolocation signals is concentrated around 4.5 kHz thus, we further investigated whether hearing loss specific to the signals' core frequency range was sufficient to impair performance. Bilateral and unilateral hearing loss for 3.5 – 5.5 kHz, but not 1.5 – 3.5 or 5.5 – 7.5 kHz, resulted in a significant drop in performance for the echo-localisation task, but not the echo-detection task. We then tested whether these findings generalized to a richer acoustic environment containing dynamic cues as well, i.e. performance in a virtual echo-acoustic navigation paradigm, as this would be more representative of what happens in real-world echolocation scenarios. We found that there were no impairments in participants' performance in this context. Our results suggest that performance was surprisingly robust to simulated hearing loss, suggesting that visually impaired people may still benefit from using echolocation despite hearing loss.

P2.3. Measuring the effect of visual cueing and beamformer steering angle accuracy on speech intelligibility in babble using a listener-in-the-loop paradigm in virtual reality

Alastair H. Moore[1], Patrick A. Naylor[1], Miike Brookes[1] & Tim Green[2]

[1]Imperial College (London, UK); [2]UCL (London, UK)

Understanding speech in a background of babble noise is challenging for human listeners and existing speech enhancement algorithms find it difficult to distinguish between the target talker and the background, particularly at low signal-to-noise ratios (SNRs). Beamforming preferentially enhances sounds arriving from a particular direction of arrival (DoA), therefore increasing the SNR, provided the selected direction co-incides with the target. DoA estimation has been extensively studied but a limited amount of attention has been given to the context of head-worn microphone arrays, such as hearing aids, hearables and smart glasses, where rapid head rotations are common.

In this study we use our recently-developed virtual reality platform for conducting listener-in-the-loop audiovisual speech intelligibility tests to investigate the interaction between beamforming strategies and head turning behaviour. Where the steering direction is fixed with respect to the head, visually cueing the target direction together with orienting towards it substantially improves speech intelligibility. However, adaptively steering the beam towards the true target direction gives even better performance and is independent of any visual cues or head turning behaviour. Adaptively steering the beam based on an estimated target direction is less effective than a fixed beamformer, even without visual cues. The results suggest there is a worthwhile benefit to be gained from more accurate DoA estimation.

Session 3: The Physiology of Hearing

Chair: Kerry Walker (Oxford)

Oral presentations

O3.1 Visual signals in ferret auditory cortex

Rebecca Norris[1], Stephen Town[1], Katherine Wood[2] & Jennifer Bizley[1]

[1]University College London, UK; [2]University of Pennsylvania

Multisensory integration is a fundamental property of mammalian sensory systems, allowing the brain to combine information across sensory modalities to parse and respond to a complex environment. Audio-visual integration is of particular interest to understanding auditory processing, as visual information influences auditory scene analysis, multi-modal object formation and speech processing. We now know that cross-modal integration happens at many stages of the processing hierarchy, including primary sensory cortices once thought to be unimodal. We have previously demonstrated visual signals are present across both primary and secondary fields of ferret auditory cortex, and our findings suggest a substantial amount of these signals enter auditory cortex via connections from visual areas suprasylvian cortex and area 21. Here, we present recent work to elucidate which features of visual stimuli are most effective for driving audio-visual integration in auditory cortex in awake, passively listening animals.

O3.2 Epoch dependent encoding of category, decision, and reward in localized regions of ferret frontal cortex

Jeffrey Boucher, Shihab Shamma & Yves Boubenec

ENS (Paris, France)

Multiple regions of frontal cortex have been shown to encode a variety of task-related variables, possibly with major inter-species differences (Mante et al, 2013, Rogers & DeWeese 2014, Siegel et al, 2015, Fritz et al 2010, Lui et al 2020, Reinert et al 2021). Though there is much overlap in the cellular response profiles in these regions, with a sufficiently complicated task and a proper methodology, key differences and interactions may be mapped. To this purpose, we trained four ferrets on a Go/NoGo auditory task, alternating in blocks which each required flexible attention to different acoustic dimensions. Using functional Ultrasound Imaging (fUS), we recorded hemodynamic responses in multiple functional areas of ferret frontal cortex while the animals performed the alternating task. We were able to observe consistent, behavioral category (Go/NoGo) dependent responses in premotor cortex (PMC) across four ferrets from stimulus presentation into the delay period. The pattern of category encoding shifted in a consistent manner between stimulus, delay, and reward periods in a manner that depended on the distance of the stimulus from the learned category boundary. In addition, consistent, well-localized categorical responses were observed in orbital cortex beneath the PMC. These results elucidate the manner in which these different functional regions represent variables in service of accomplishing the animals' goals, and suggest targets for further experimentation.

O3.3 Genetic risk for schizophrenia and experience of hearing impairment both influence auditory brain function in a mouse model of 22q11.2 Deletion Syndrome

Chen Lu & Jennifer F. Linden

UCL Ear Institute, London, United Kingdom

Hearing impairment has been identified as a longitudinal risk factor for schizophrenia (Linszen et al. 2016 Neurosci Biobehav Rev), but the causation for this association are not well understood. The 22q11.2 chromosomal microdeletion is one of the strongest known genetic risk factors for schizophrenia, and up to 60% of carriers have mild to moderate hearing impairment (Verheij et al. 2017 Clin Otolaryngol 2017). Here we used Df1/+ mouse model of 22q11.2 Deletion Syndrome (22q11.2DS) to investigate how hearing impairment interacts with genetic risk for schizophrenia to affect auditory brain function. The Df1/+ mouse replicates the large inter-individual variation in hearing ability observed among 22q11.2DS patients and exhibits auditory brain abnormalities consistent with disrupted cortical excitation/inhibition balance (Zinnamon et al. 2022 Biol Psychiatry Global Open Sci). We measured peripheral hearing sensitivity and cortical auditory evoked potentials (AEPs) in 29 Df1/+ mice and 22 WT littermates, exploiting inter-individual variation in hearing ability to distinguish effects of hearing impairment from those of genetic risk for schizophrenia. AEP measures of cortical gain and adaptation were abnormal in Df1/+ mice with normal hearing, but were also affected by hearing impairment. Our results show that auditory cortical abnormalities in 22q11.2DS depend on both the genetic deletion and experience of hearing impairment.

O3.4 Environmental noise modulates sound encoding in adult rat auditory cortical neurons

Natsumi Homma[1], Natsumi Homma[2] & Christoph Schreiner[2]

[1]University of Cambridge (Cambridge, UK); [2]University of California San Francisco

Cortical sound representation can be altered in response to environmental sound statistics and behavioral task engagement. This cortical plasticity is strongest during the critical period; however, the ability is somehow maintained throughout the adulthood. Here we investigated if adult animals have comparable plasticity to the critical period and how sound encoding in auditory cortical neurons changes by altering sound statistics in the environment. We exposed adult Sprague-Dawley rats to spectrotemporally modulated noise at a moderate level (~60 dB SPL) for eight weeks and, in a separate exposure group, provided three-weeks quiet period after the noise exposure. Then, we investigated neural responses in the primary auditory cortex (A1) and ventral auditory field to investigate the ability of signal-in-noise processing and spectrotemporal properties. We found that adult noise-exposed animals improved cortical signal encoding compared to unexposed control animals for both groups. Moreover, spectral receptive fields in A1 shifted away from the exposed noise statistics, which were similar to those observed in the critical-period exposed animals. By contrast, noise-exposed animals followed by a quiet period showed plasticity effects more related to temporal processing. These findings support that environmental sound statistics can affect hearing ability even after the critical period and indicate potential implications for hearing protection and approaches to hearing re-training

Poster presentations

P3.1 Flexible integration of natural stimuli by auditory cortical neurons

Grace Ang, Claudia Clopath & Andriy Kozlov

Imperial College London

While many experiments characterise the neuronal input-output function by directly activating synaptic inputs on dendrites in vitro, the integration of real-world stimuli presenting complex spatiotemporal inputs to the neuron is less well studied. Using ethologically relevant stimuli, we study neuronal integration in relation to Boolean functions thought to be important for pattern recognition. In the mouse auditory cortex, we record single-unit responses to pairs of ultrasonic mouse vocalisation (USV) syllables. We observed a range of integration responses, spanning the sublinear to supralinear regimes, with many responses resembling the MAX-like operation. Integration was more MAX-like for strongly activating features, and more AND-like for spectrally distinct inputs. Importantly, single neurons could implement more than one integration function, in contrast to artificial networks which fix activation functions across all units and inputs. To understand the mechanism underlying the flexibility and heterogeneity in neuronal integration, we modelled how dendritic properties could influence the integration of inputs with complex spectrotemporal structure. The results link nonlinear integration in dendrites to single-neuron computations for object recognition.

P3.2 Neural mechanisms of stream formation during active listening in the ferret auditory cortex

Jules Lebert[1], Carla Griffiths[1], Joseph Sollini[2] & Jennifer Bizley[1]

[1]*University College London;* [2]*University of Nottingham*

In real-world listening, we interpret mixtures of overlapping sounds. The brain breaks down these scenes into individual objects, forming auditory streams. We study how the auditory cortex contributes to stream formation. The temporal coherence theory explains stream formation by grouping sounds based on their coherence along different feature axes. However, this theory hasn't been tested with naturalistic sounds at the neural population level. We trained ferrets to detect target words in streams of random distractor words, amidst spatially separated noise. Neural data from ferret auditory cortex revealed that word encoding is unaffected by stream location in a single-stream task. However, in the presence of a competing stream, word encoding in the contralateral stream outperformed the ipsilateral stream. Early in the trial, decoding of the ipsilateral stream was at chance level, but it improved towards the end, approaching the contralateral stream. These findings suggest that auditory cortical populations quickly adapt to switch from representing the contralateral stream to the attended stream in the presence of competing noise.

P3.3 Investigating how neurons invariantly encode pitch derived from two types of acoustic cues

Veronica Tarka[1], Quentin Gaucher[2] & Kerry Walker[1]

[1]*Oxford University, Oxford;* [2]*Paris Saclay Institute of Neuroscience*

Pitch is our perception of the tonal quality of sound. It is the basis of musical melody and plays a key role in communication and sound segmentation. The pitch is perceived at a single fundamental frequency (F0), which can be derived for a complex sound from either the regular spacing of harmonics in the frequency domain, the repetition rate of the sound's waveform in time, or a combination of these features. Studies in marmosets have described specialized neurons located within a "pitch centre" in auditory cortex that encode F0 invariantly to other spectral changes, but there has not been clear evidence for such specialized pitch neurons in other species. We performed Neuropixels recordings of single neurons in the auditory cortex of 4 anaesthetized ferrets while presenting a variety of pitch-evoking sounds across a range of F0s (0.25-4 kHz). We found that some neurons derived F0 exclusively from resolved harmonics, while others from temporal periodicity. A further subset of neurons encoded F0 invariantly across both classes of pitch cues, which may be the first evidence for specialized "pitch neurons" in non-primates. These neurons were not confined to a localized pitch centre, but were instead distributed throughout primary auditory cortex (A1 and AAF). F0 tuning in these neurons was robust across many complex sounds, but usually failed to extend to the frequency of pure tones, suggesting they may be specialized to represent the pitch of complex sounds.

P3.4 Investigating the causal role of auditory cortex in auditory processing and adaptation to noise

Louay Madanat[1], Maya Khalil[1], Katarina C. Poole[2] & Jennifer Bizley[1]

[1]UCL Ear Institute (London, UK); [2]Imperial College London (London, UK)

The brain's construction of the auditory world demands complex processing in multiple brain regions. These regions serve multiple functions, likely operating in dynamic processing networks. The auditory cortex (AC) is a critical part of this process, with primary auditory cortex being required for accurate sound localisation and for discriminating speech sounds in noise. Human work suggests that secondary AC areas may be functionally specialised and play a key role in adaptation to noise. This project's aim is to investigate the role of non-primary AC by using optogenetic methods to reversibly inactivate AC subfields in ferrets during the performance of complex listening tasks. To refine these tasks, work in humans is mapping the timecourse of adaptation to background noise statistics in sound localisation and speech-in-noise tasks. Data collection is ongoing but preliminary analysis suggest that this process is very rapid (on the order of ~100-300 ms). We have used stimulation of inhibitory interneurons via mDlx-directed Channelrhodopsin-2 (ChR2) to silence auditory cortex in ferrets with the ultimate goal of assessing the role of auditory cortex in adaptation to background noise as well as target identification and localisation.

P3.5 Plasticity of temporal integration in ferret auditory cortex

Magdalena Sabat[1], Quentin Gaucher[1], Sam Norman-Haignere[2] & Yves Boubenec[1]

[1]École Normale Supérieure, PSL University, CNRS, 75005 Paris, France; [2]University of Rochester Medical Center, Rochester, NY, USA

Retrieving meaning from complex auditory stimuli requires flexible binding in time of auditory features at different timescales. Recent reports suggest that auditory perception relies on a hierarchy of auditory timescales and progressive processing of information through a hierarchy of cortical areas. Although temporal dynamics of auditory perception have been thoroughly studied, little is known about the plasticity of temporal integration windows (or timescales) in the auditory cortex. To address this question we study the neural responses of the auditory cortex of the ferret to complex continuous auditory stimuli during behavior. First, we assess the similarity of the gradient of sensory timescales in the auditory cortex of the ferret to that of the human. Subsequently, we address the question of plasticity by comparing the structure of this gradient between passive exposure and active behavior. Finally, we explore other sources of variability such as the internal state of the animal by relating the structure of the gradient to the physiological state of the animal. This work provides insights into the cortical mechanisms underlying flexible auditory performance at multiple scales.

P3.6 Auditory cortex spontaneously reactivates recent sensory experiences

Arsenii Goriachenkov & Yves Boubenec

Ecole Normale Superieure

Intense dialogue between the hippocampus (HC) and cortical structures underlies memory consolidation. Spontaneous reactivations, which usually occur during slow-wave sleep or calm awake state, mediate the transfer of past experience from HC to the cortex for long-term storage. Reactivations have often been observed in HC, but rarely in the cortex.

To investigate whether sensory experience elicits spontaneous reactivations in the auditory cortex (AC), we recorded neural activity different regions of the AC in the awake head-fixed ferret using functional UltraSound (fUS) neuroimaging before, during and after passive exposure to ferret vocalizations, music, and human speech. Using the classical method of identification of co-firing neuronal assemblies [Peyrache, 2010], we were able to demonstrate the existence of large-scale cortical reactivations in awake animals.

We also show that the behavioural relevance and the novelty of the presented sound are correlated with the number and the magnitude of its subsequent reactivations. To test the effect of novelty, we compared two recording sessions held with an 8-hour delay. We observed significantly more reactivations during the first exposure to the sound than during the second exposure (panel B). This effect was more important for species-relevant sounds (ferret vocalizations) rather than for human speech or music (panel C).

These results confirm that reactivations of sensory experience are indeed present in spontaneous cortical activity. Moreover, the magnitude of reactivations is increased by the novelty and the relevance of the experience.

P3.7 Effects of sensorimotor feedback in ferret auditory cortex during acoustic production

Flavien Féral, Félicie levinton & Yves Boubenec

Ecole Normale Superieure

This study investigates the intricate relationship between action and perception that underpins numerous behaviors, specifically those involved in acoustic signal production. The focus lies on the coordination of motor and sensory neural processes, which are crucial for the precise control of complex sounds found in human speech or animal vocalizations.

Our objective was to examine the role of sensorimotor interactions and the impacts of sensorimotor task training on the signals in the auditory cortex. We developed a device synchronizing the animal's head movements with the frequency of an acoustic feedback, generating self-produced sound through a closed-loop feedback.

Two conditions were contrasted: 'tracking,' where sound frequency was linked to the ferret's head position, and 'playback,' where this correlation was absent. A significant decrease in overall neural response during tracking (closed-loop), compared to playback (open-loop), was discerned in both primary and secondary auditory cortices. Interestingly, suppressive effects were stronger in primary, than non-primary regions of the auditory cortex. Neural response to tones matched between tracking and playback during motion reduced drastically, while non-matching tones exhibited modest suppression, suggesting an online encoding of the predictability of the incoming sensory inputs.

Our findings provide compelling evidence suggesting that sensorimotor interactions play a significant role in modifying neural signals in the auditory cortex. The diminished response during tracking suggests that there is an interplay between expected auditory signals and actual sensory outcomes that is central to motor control and learning.

P3.8 Age, high-frequency hearing function and speech perception in real-world settings: a cross-sectional study

Ben Lineton, Nicci Campbell, Mashael Alshayie & Khimu Pun

ISVR, University of Southampton

Recent studies suggest that access to high-frequency components of speech (>8 kHz) can be important for speech perception in some challenging situation (Hunter et al, 2020). Results will be presented from a cross-sectional study of forty participants with normal audiometric thresholds from 0.25 to 2 kHz, and who fall into two ages groups: 18-40 and 50-70 years. Measurements of speech perception in noise tests will be made of speech intelligibility with roving speech-in-noise using the crescent-of-sound rig with speech material recorded to capture high-frequency components up to 20 kHz. Assessments will also be made of extended-high frequency hearing thresholds (up to 16 kHz), distortion product otoacoustic emissions (up to 10 kHz) and self-reported hearing. A correlational analysis of results will be presented.

Hunter, L. L., Monson, B. B., Moore, D. R., Dhar, S., Wright, B. A., Munro, K. J., Zadeh, L.M., Blankenship, C. M., Stiepan, S. M., and Siegel, J. H. (2020). “Extended high frequency hearing and speech perception implications in adults and children” Hearing research, 397, 107922.

P3.9 The effect of cochlear transducer operating point on stimulus-frequency otoacoustic emissions derived from a nonlinear cochlear model

Vaclav Vencovsky & Ales Vetesnik

Czech Technical University in Prague, Czech Republic

Conductance of mechanoelectrical transduction channels in the hair bundle of outer-hair cells changes nonlinearly with the hair bundle displacement. In a cochlear model, we can simulate the nonlinear function with a Boltzmann function. Here we examine the effect of operating point of the nonlinearity on the stimulus-frequency otoacoustic emissions (SFOAEs) derived from a hydrodynamical cochlear model. In our previous work with the same model, we found a semi-analytical solution for SFOAEs. The solution has two components: a component due to linear reflection from mechanical irregularities and a component due to perturbation of the nonlinear force. The nonlinear force is perturbed by the reverse traveling wavelets which are reflected from mechanical irregularities. A change in the operating point of the nonlinear Boltzmann function affects the intensity growth of the nonlinear component of SFOAEs. Because the reflection component and the component due to perturbation of the nonlinear force have opposite phases, they destructively interfere. This interference affects the input/output function of SFOAEs.

P3.10 Effects of Stimulus Polarity and Gaze Position on Ocular Vestibular Evoked Myogenic Potentials (oVEMPs) Responses Using B71 and Minishaker Transducers

Busra Kocak Erdem, Steve Bell & Ying Ye

University of Southampton

The Vestibular Evoked Myogenic Potentials (VEMPs) are used to assess the vestibular system by examining the otolith organs. Ocular VEMPs are responses recorded from contralateral extraocular muscles. Our objectives were to investigate the impact of stimulus polarity and gaze position on oVEMPs using B71 and minishaker transducers. Ten participants, each with normal hearing and without any back, neck, or balance problems, were assessed to stimulation at four different frequencies (125, 250, 500, and 750 Hz) to investigate the effect of polarity changes. And also, oVEMPs were evaluated at 125 Hz under both straight and upward gaze conditions. Results showed that high-frequency (500 and 750 Hz) oVEMP responses using the B71 were unaffected by polarity changes, with standard responses observed, whereas low-frequency responses (125 and 250 Hz) were affected, particularly with the minishaker ($p<0.05$). The minishaker and B71 were affected by the gaze position at 125 Hz with maximum output ($p<0.05$.) Low-frequency responses were found to be sensitive to stimulus polarity changes and were enhanced by single polarity stimulation with the minishaker. The upward gaze position also increased response amplitude. At low frequencies, more output was achieved with the minishaker, but at higher frequencies (500 Hz), the transducers had similar outputs. Thus, the B71 is not able to stimulate low frequencies in the way that the minishaker can.

P3.11 Optimization of the measurement method for cochlear transducer function by means of nonlinear hydrodynamic cochlear model

Aleš Vetešník & Václav Vencovský

Czech Technical University in Prague, Czech Republic

Meniere's disease is accompanied with endolymphatic hydrops, which, can cause a change in the gain of the cochlear amplifier as well as a change in the position of its operating point. Otoacoustic emissions (OAEs) might provide a non-invasive objective insight into the activity of the cochlear amplifier. Distortion product OAEs (DPOAEs) were used to derive the nonlinear function of the cochlear amplifier. In particular, the method uses a low-frequency bias tone to shift the two-tone signal through the nonlinear function which leads to the modulation pattern of DPOAE amplitude. This pattern approximates the third derivative of the nonlinear function provided that the levels of primary tones

are sufficiently small and that DPOAEs are generated only in a narrow region of the basilar membrane, so that it cannot lead to the destructive interference among elementary DPOAE sources.

We used a nonlinear hydrodynamic model of the cochlea to verify for which input amplitudes of primary tones these conditions are met. The model allows for using a semi-analytical solution to analyze how individual elementary sources on the basilar membrane contribute to the overall DPOAE. Based on the simulations for various amplifier gain, we can propose optimal conditions for obtaining the parameters of the nonlinear function of the cochlear amplifier.

Session 4: Objective Measures of Hearing (EEG, fNIRS, fMRI etc.)

Chair: Chris Sumner (Nottingham Trent)

Oral presentations

O4.1 Studying the neural basis of changes in the perceived loudness following the activation of a cochlear implant

Dorothée Arzounian[1], François Guérit[1], John M. Deeks[1], Charlotte Garcia[1], Evelien de Groote[1], Manohar Bance[2] & Robert P. Carlyon[1]

[1]Cambridge Hearing Group, MRC Cognition and Brain Sciences Unit, University of Cambridge (Cambridge, United Kingdom); [2]Department of Clinical Neuroscience, University of Cambridge, Cambridge, United Kingdom

The electrical stimulation levels needed to produce a comfortably loud sensation usually increase over the first 6 months following the activation of a cochlear implant (CI), but the neural basis for this effect remains unknown. Here we present a longitudinal study of the changes in neural responses to electrical stimulation that accompany these changes in stimulation levels. We combine neural response telemetry and electro-encephalography to measure a set of electrically-evoked responses known to have generators at different levels of the central auditory pathway, from auditory nerve to cortex: the Compound Action Potential, Auditory Brainstem Response and Auditory Steady-State Response. Measures are obtained at 4 time points between CI activation and 4 months later. Whenever possible, we compare amplitudes of responses obtained at different time points using the same current level in order to rule out or reveal evidence for long-term adaptation. This is done at each central level probed. We also compare, between time points, responses to stimulus levels yielding a constant, comfortable perceived loudness. At the time of abstract submission, data collection is ongoing. We expect complete datasets for 5 participants and partial data for 4 more at the time of the meeting. Early results suggest that, even at the highest central levels tested, neural adaptation only partially accounts for the increase in stimulation level required to maintain perceived loudness.

O4.2 Electrophysiological evidence for prediction errors during perception of degraded spoken sentences

Ediz Sohoglu & James Webb

University of Sussex

Prediction facilitates language comprehension but how are predictions combined with sensory input during perception? Previous work suggests that cortical speech representations are best explained by prediction error computations rather than the alternative ‘sharpened signal’ account (Blank & Davis 2016 PloS Biol). However, this previous work used an artificial listening situation (single word listening and predictions obtained from external written cues). In the current work we explore a more naturalistic listening situation in which listeners heard sentences and predictions obtained directly from the speech signal i.e. from sentence context.

Listeners (N=30) heard degraded (16-channel noise-vocoded) sentences in which the last word was strongly or weakly predicted by the preceding context, based on cloze probability. All sentences were semantically coherent. We also manipulated signal quality of the final word (2/4/8 channels). Using Temporal Response Function (TRF) analysis of EEG responses to the final word, we measured cortical representations of speech acoustic features (spectral and temporal modulations). We observed a significant interaction between prediction strength and signal quality (p = .02) such that TRF model accuracies increased with strong predictions but only when signal quality was low. These findings are more consistent with the prediction error account (Blank & Davis 2016) and show that previous findings extend to more naturalistic listening situations.

O4.3 Frequency mapping in human primary auditory cortex predicts frequency discrimination performance

Julien Besle[1], Benjamin Gurer[2], Rosa-Maria Sanchez-Panchuelo[2], Susan Francis[2] & Katrin Krumbholz[2]

[1]University of Plymouth, UK; [2]University of Nottingham, UK

Cochlear characteristic frequency increases approximately exponentially with basilar membrane position, but it is unclear whether the same frequency mapping is preserved in auditory cortex. In the visual system, foveal magnification is greater in primary cortex than in the retina, and predicts positional acuity better. Here, we measured the frequency mapping function in primary auditory cortex using 7-Tesla fMRI and compared it with predictions from published estimates of human cochlear frequency mapping and frequency discrimination performance.

We measured BOLD frequency tuning curves between 0.1 and 8 kHz in 19 normal-hearing listeners (sparse GE-EPI, TR = 7.5s, TA = 2.2s, TE = 25 ms, FA = 90º), and estimated voxelwise preferred frequency and frequency selectivity by fitting Gaussians in log-frequency space. All participants showed two mirror-reversed tonotopic gradients separated by a low-frequency reversal along Heschl's gyrus, in a region of elevated frequency selectivity corresponding to primary auditory cortex.

In both tonotopic gradients, cortical distance from the low frequency reversal increased fastest with preferred frequency around 1-2 kHz, indicating maximal cortical magnification at these frequencies. This frequency mapping is better predicted by frequency discrimination performance than by the cochlear frequency mapping function. This could indicate that human frequency discrimination performance is constrained by cortical and not cochlear magnification.

O4.4 Gaussian Processes for efficient audiogram estimation with Auditory Brainstem Responses

Steven Bell, Michael Chesnaye, Josef Schlittenlacher & David Simpson

ISVR, Southampton

The Auditory Brainstem Response (ABR) is used to determine hearing thresholds in newborns. Our objective was to reduce test time for ABR audiogram estimation using a Gaussian Process (GP): a Bayesian approach for non-linear regression. This is combined with active learning methods to efficiently sample the frequency-intensity input space and maximize the amount of information gained within limited test time. Audiograms were estimated to band-limited chirp stimuli at 4 frequencies in 22 normal-hearing and 9 hearing-impaired subjects using 1. a behavioural approach 2. visual inspection of ABRs by examiners 3. A GP approach applied to ABR. The GP estimates ABR peak-to-peak amplitudes as a function of intensity, from which thresholds are derived. Bootstrap estimates of ABR amplitudes were used as GP inputs. Comparisons were drawn between the GP and the standard BSA approach in terms of test time and accuracy. Average test times for audiogram estimation were 31.2 minutes for the GP, and 57.1 minutes for visual inspection. Mean hearing threshold estimation errors were 0.5 dB for the GP and 7.2 dB for visual inspection. The GP reduced test time by approximately 45% compared to visual inspection by clinicians. Average hearing threshold estimation error was also lower for the GP, but the errors were more variable (and occasionally quite high), although there is scope to reduce this by optimising the threshold algorithm choice. The GP has potential for reducing ABR measurement time.

Poster presentations

P4.1 Assessing array-type differences in current spread in cochlear implant users using the Panoramic ECAP Method

Robert Carlyon & Charlotte Garcia

University of Cambridge, UK

The amount of current spread from cochlear implant (CI) electrodes and the status of surviving neurons varies between patients and along the length of the electrode array in individual patients. These variations may help explain why some listeners struggle to hear clearly & may guide patient-specific programming methods. The Panoramic ECAP (PECAP) method separately estimates current spread and neural responsiveness from ECAPs obtained with all possible combinations of masker & probe electrodes (Garcia et al, 2021). Here we present PECAPs from a large number of CI users with different array types, predicted to differ in the proximity of electrodes to auditory neurons. This provides normative data for clinicians and tests the assumption that the amount of current spread will differ between array types.

PECAPs were analysed for 90 users of Cochlear © CIs with perimodiolar, slim-straight, and slim-modiolar arrays (n=43, 25, & 12). We found highly significant main effects of array type and electrode on current spread, as well as a highly significant interaction. Slim-modiolar arrays showed the lowest current spread, followed by the perimodiolar arrays, and slim-straight arrays. Apical electrodes showed higher current spread than the rest of the array. Slim-straight arrays showed wider current spread than the others did at the apex, specifically. A significant but much smaller effect of array type was found on the normalised estimate of neural responsiveness.

P4.2 How the brain tracks structure in rapid and slow sound sequences – a MEG study in humans

Mingyue Hu[1], Antonio Rodriguez Hidalgo[1], Roberta Bianco[2] & Maria Chait[1]

[1]Ear Institute, UCL, London, UK; [2]Neuroscience of Perception & Action Lab, Italian Institute of Technology, Rome, Italy)

Human listeners are sensitive to patterns in rapidly unfolding sound sequences. This ability can be measured with MEG/EEG, revealing fast memory formation of transition probabilities within the stimulus sequence. In this MEG study, we asked if listeners can track repeating patterns in slow tone sequences and which brain networks are involved. Stimuli consisted of 50ms tones arranged in random (RND) or regularly repeating patterns (REG). Two timing profiles were used: 'fast' sequences consisted of contiguous tones (20Hz presentation rate); 'slow' sequences had tones separated by a 200 ms silent gap (4Hz rate). Participants listened passively to sound sequences. Passively elicited MEG brain responses showed significantly stronger sustained response magnitude in REG relative to RND. This was observed even in the slower sequences, despite the long durations of the pattern (2500ms), revealing the auditory brain's remarkable implicit sensitivity to complex patterns. Importantly, brain responses evoked by single tones exhibited the opposite pattern - stronger responses to tones in RND compared to REG sequences. The overall response pattern is not consistent with increased gain on predictable sensory information (e.g. as hypothesized by predictive coding) but is in-line with increased inhibitory activity in REG sequences. The observation of simultaneous but opposing DC and evoked response effects reveal concurrent processes that shape the representation of unfolding auditory patterns.

P4.3 The effect of brief interruptions on the neural representation of ongoing scene statistics: An EEG study

Kaho Magami & Maria Chait

Ear Institute, UCL (London, UK)

The auditory system constantly tracks environmental statistics ('context') and predicts upcoming events, even when not consciously attending to sound. However, how and whether the old context is retained in memory following scene interruptions remains unknown. This study aimed to examine the effect of interruption duration on context-relearning efficiency. How is ongoing context representation affected by an interruption, such as a passing train? Thirty-one participants passively listened to a rapidly evolving regular sound pattern while EEG responses were measured. The 75% of scenes included a brief interruption (1, 3, or 5 tones) partway. We show that the learning trajectory of the sound context is reflected in the EEG sustained response. Power increased with exposure to the sound context, dropped upon interruption onset, and recovered as the original context re-emerged. Notably, the recovery slope varied across conditions, with shorter interruptions yielding steeper slopes. These dynamics are consistent with predictions from an ideal observer model, which quantifies the predictability of each tone-pip in the sound sequence. In summary, this study demonstrated that the speed to relearn the previously encountered context depends on the duration of the interruption signal. This finding also suggests that the brain does not disregard even brief interruptions as mere noise, necessitating relearning of the immediately preceding context.

P4.4 How underlying statistical structures modulate the neural response to rapid auditory sequences

Alice Milne & Maria Chait

University College London

The brain is highly sensitive to auditory regularities and this is exploited in many scenarios from parsing complex auditory scenes, to language acquisition. To understand the impact of stimulus predictability on perception, it is important to determine how the detection of a predictable structure influences processing and attention. Here we probed how the brain response differs based on the predictability of an auditory sequence. Using an EEG paradigm we tested the neural response to sequences of 50ms tones arranged into a random order, a deterministic pattern or a probabilistic structure where the transitional probabilities between tones allowed them to be segmented into triplets. In addition, we introduced deviant tones that were outside the spectral frequency of the main sequence, predicting based on previous evidence, that there would be a stronger deviant response in more predictable sequences.

We found that the brain rapidly detects the underlying structure, locking to the rate of the triplets. Furthermore, the sustained neural response is modulated by different forms of predictability. Finally, we demonstrate that the event-related response to deviant tones is influenced by both sequence type and the position of the deviant in the triplet structure. We discuss our findings in relation to cognitive resource allocation and the predictive coding framework.

P4.5 Comparison of auditory cortical N1 responses evoked by regularity detection in different auditory contexts

Buse Adam[1] & Mehmet Yarali[2]

[1]*University College of London;* [2]*Hacettepe University, Turkey*

Understanding the mechanisms of the hearing system requires investigating the detection various types of auditory regularities within different background contexts. In this research, we aimed to compare the auditory cortical N1 responses evoked by regular sound patterns with a decreasing frequency which is embedded in random sounds (rand_dec_2) and regular sound patterns with decreasing frequency in isolation (rand_dec), both appearing after random frequency sounds. It is noteworthy that the decreasing frequency pattern used in our study may resemble intonation patterns and Fo trajectories commonly found in everyday speech, such as while making a statement. Therefore, using different types of regularity could be an important point of our work. A total of 20 young participants were recruited for the study. Behavioural results indicated that regularity detection for rand_dec_2 was significantly more challenging compared to rand_dec. Furthermore, both types of regularities elicited N1 responses, albeit with significantly smaller amplitudes and longer latencies observed for the rand_dec_2. In addition, there were significant negative correlations between the N1 amplitude values and hit rates for rand_dec_2. Overall, the findings show that different types of auditory regularities can also be detected when they are embedded in unrelated sounds, and they have cortical representations.

P4.6 The Relationship Between eCAP Latency and Characteristic Frequency of Human Spiral Ganglion Cells

Jason Lien & Debi Vickers

University of Cambridge

Introduction

Spiral Ganglion Cells (SGCs) with greater numbers of potassium channels can recover from firing faster and hence capable of phase-locking to a higher frequency. Rodent studies (Adamson et al., 2002; Liu & Davis, 2007) using patch-clamp techniques reported that SGCs with higher Characteristic Frequencies (CF) have more potassium channels and it's reflected on the electrophysiological parameters recorded such as the stimulus to spike latency and action potential (AP) duration.

It is hypothesised that human SGCs with higher CFs have shorter AP durations.

Method

Data Collection

Electrically evoked compound action potential (eCAP) data measured in different sites of cochlea from 11 Cochlear and 16 MED-EL users were shared.

Pre-processing

Raw eCAP recordings were fitted a curve to facilitate automatic N1, P2 labelling. The site of stimulating electrode was transformed into insertion angle based on the average values for the type of electrode (Landsberger et al., 2015). Analysis on the recovery function (recovery inter-pulse-interval) is ongoing and findings will be presented.

Preliminary Results

A linear regression was conducted for each participant to relate estimated insertion angle to eCAP P2-N1 latencies, creating 27 coefficients. A two-tailed 1-sample t-test reported that the 27 coefficients are highly unlikely to come from a distribution with the mean of 0 ($t = 4.12$, $p = 0.003$, $df = 26$). Thus, there may be a relationship between insertion depth and eCAP P2-N1 latency.

Session 5: Psychoacoustics and Cognitive Aspects of Normal and Impaired Hearing

Chair: Brian Moore (Cambridge)

Oral presentations

O5.1 Pupil Dilation and Microsaccades Provide Complementary Insights into the Dynamics of Arousal and Instantaneous Attention during Effortful Listening

Maria Chait, Claudia Contadini-Wright, Kaho Magami, Miaohsu Chang, Xinping Fu & Nischay Mehta

Ear Institute, University College London, London, UK

Listening in noisy environments involves effort and arousal. Separately quantifying these components helps uncover the dynamics of effort and how it changes across situations. We concurrently measured two types of ocular data in young participants: pupil dilation (PD; thought to index arousal aspects of effort) and microsaccades (MS; hypothesized to reflect automatic visual exploratory sampling), during a speech-in-noise task under high (HL) and low (LL) listening load conditions. Sentences were manipulated so behaviorally relevant information appeared at the end (Experiment 1) or beginning (Experiment 2) of the sentence, resulting in different temporal demands on focused attention. We observed a sustained difference in PD between HL and LL conditions consistent with increased phasic and tonic arousal. MS rate was also modulated by listening load, seen as a reduced MS rate in HL relative to LL. Critically, MS effects were localized in time during periods when demands on auditory attention were greatest. This suggests that auditory selective attention interfaces with MS-generating mechanisms, establishing MS as an informative way to quantify the temporal dynamics of auditory attentional processing in effortful listening. Robust follow-up data collected on a portable eyetracker demonstrates paradigm adaptability beyond the laboratory setting. Comparable results in older adults suggest PD and MS can be reliably used to probe arousal and attentional dynamics in aging populations.

O5.2 Relative pitch representations and invariance to timbre

Malinda J. McPherson[1] & Josh H. McDermott[2]

[1]*University of California, San Diego, (La Jolla, USA);* [2]*MIT*

Information in speech and music is often conveyed through changes in fundamental frequency (f0), perceived by humans as “relative pitch”. Relative pitch judgments are complicated by two facts. First, sounds can simultaneously vary in timbre due to filtering imposed by a vocal tract or instrument body. Second, relative pitch can be extracted in two ways: by measuring changes in constituent frequency components from one sound to another, or by estimating the f0 of each sound and comparing the estimates. We examined the effects of timbral differences on relative pitch judgments, and whether any invariance to timbre depends on whether judgments are based on constituent frequencies or their f0. Listeners performed up/down and interval discrimination tasks with pairs of spoken vowels, instrument notes, or synthetic tones, synthesized to be either harmonic or inharmonic. Inharmonic

sounds lack a well-defined f0, such that relative pitch must be extracted from changes in individual frequencies. Pitch judgments were less accurate when vowels/instruments were different compared to when they were the same and were biased by the associated timbre differences. However, this bias was similar for harmonic and inharmonic sounds and was observed even in conditions where judgments of harmonic sounds were based on f0 representations. Relative pitch judgments are thus not invariant to timbre, even when timbral variation is naturalistic, and when such judgments are based on representations of f0.

O5.3 Effects of hearing loss on listener's ability to detect mistuning in music

Sara Miay Kim Madsen[1], Heather A. Kreft[2], Elaea Purmalietis[2], Torsten Dau[3] & Andrew J. Oxenham[2]

[1]Technical University of Denmark (Lyngby); [2]UMN; [3]DTU

This study investigated the effect of hearing loss on mistuning detection in music, and on the detection of two potential mistuning cues of envelope fluctuations (beats) and inharmonicity. We tested 30 listeners with sensorineural hearing impairment (HI), 30 older listeners with audiologically normal hearing (ONH), and 31 younger normal-hearing listeners (YNH). Listeners judged whether short excerpts from J.S. Bach's Two-Part Inventions were in tune. We tested mistuning detection for varying mistuning sizes (experiment 1) and for conditions that retained or limited either beating or inharmonicity cues (experiment 2). We also tested detection thresholds for inharmonicity and beats for diagnostic pure-tone and complex-tone stimulus, and frequency selectivity via an abbreviated notched-noise method. Mistuning detection was significantly poorer for the HI group than the other two groups (experiment 1), but no significant between-group differences were observed in experiment 2. For the diagnostic stimuli, no group differences were observed in either inharmonicity or beat detection for the complex tones, but pure-tone beat detection was better in the HI group than in the other two groups. As expected, frequency selectivity worsened with increasing hearing loss. Overall, although hearing impairment was associated with a somewhat poorer ability to detect mistuning in music, it did not affect listeners' detection or use of the underlying cues of beats or inharmonicity.

O5.4 An Association Between Auditory Responsiveness of Children and Duration of Entertainment Screen Time in the First Two Years of Life

Josephine Marriage[1], Mahmoud Keshavarzi[2] & Brian C. J. Moore[3]

[1]Chear Ltd Royston UK; [2]Centre for Neuroscience in Education, University of Cambridge; [3]Cambridge Hearing Group, University of Cambridge

There is an increase in the proportion of children, age 1-3 years, who do not respond to simple signals used in standard behavioural hearing tests. Further testing demonstrates good peripheral cochlear hearing function in these children. During the COVID pandemic young children experienced greater social isolation and spent more time viewing screens. We report factors associated with a change in observed auditory response behaviours and communication skills.

METHOD 118 children (1-3 years) attending a hearing assessment centre had behavioural hearing assessments using standard simple sounds (children assigned to Group 2) and for children who needed familiar tunes to be used (Group 1).

OUTCOME MEASURES Parents completed a questionnaire covering: the child's preferred tune; number of times the child met with other children; number of words the child spoke at one and two years; and daily amount of EST the child had over the first and second years of life. An assessment of each child's social, attention and communication skills (MoSAIC) was completed by the audiologist.

RESULTS Group 1 had significantly fewer words than group 2 and significantly higher EST at 1 and 2 years of age. Based on MoSAIC scores, group 1 showed lower social, attention and communication skills than group 2 (p = 0.0001). A large amount of EST was associated with poorer auditory responsiveness and communication skills.

Poster presentations

P5.1 Cortical Tracking of the Speech Amplitude Envelope with Applications to Cochlear Implants

Alexis Deighton MacIntyre & Tobias Goehring

University of Cambridge (Cambridge, UK)

During speech listening, patterns of neural activity become temporally coupled to stimulus features, which can be measured using electroencephalography (EEG). This "cortical tracking" may hold clinical promise as an objective measure to guide cochlear implant (CI) fitting; however, the effect of spectral degradation associated with CI signal processing on this technique is unclear. We simulate CI listening by presenting natural and spectrally degraded speech to typically hearing listeners (n = 36) undergoing EEG recording. To dissociate sensory from linguistic-phonological processing, we use intelligible (English) and non-intelligible (Dutch) speech produced by one bilingual speaker. To maintain auditory attention irrespective of intelligibility, we devised a novel prosodic target detection task. Decoding models were trained to reconstruct the speech amplitude envelope from held-out neural response data, with the correlation between deconstructed and true stimulus envelope providing a measure of cortical tracking. Cortical tracking was slightly, but significantly, reduced for non-intelligible speech. We find no clear effect of spectral degradation. The target-detection task elicited good performance across conditions, although both intelligibility and spectral degradation adversely impacted reaction times. We conclude that cortical tracking could in principle be used as an objective measure of speech envelope tracking with CI.

P5.2 Does the bilingual advantage exist when the stimuli is auditory?

Meital Avivi-Reich & Ina Selita

City University of New York (CUNY)

Some studies suggest that those who are proficient in more than a single language demonstrate better performance when requested to conduct tasks that involved Executive Functions (EF). These findings led to the term "bilingual advantage". However, there may be processing costs that come with bilingual exposure and proficiency. Speech perception studies show that bilingual listeners might experience greater difficulties when listening under adverse conditions compared with monolingual native listeners. Those greater difficulties could potentially eliminate any EF advantages when the input upon which the functions are applied is auditory. A literature review focusing on studies which tested the existence of a bilingual advantage using auditory stimuli reveals that only a limited number of such studies has been published so far using inconsistent methods. Overall, the review implies that EF tasks which rely on auditory stimuli, or a combination of visual and auditory stimuli show no bilingual advantages, while for tasks in which the stimuli is visual most of the results demonstrate a bilingual advantage. These findings imply that there may be an effect of stimuli modality on EF performance which differs between bilingual and monolingual participants. A new study will be discussed which has been designed to systematically study the effect linguistic experience may have on EF depending on the modality of the stimuli.

P5.3 Are eyes a potential objective measure of auditory scene analysis?

Mert Huviyetli & Maria Chait

UCL - Ear Institute (London/United Kingdom)

The auditory system plays a crucial role as the brain's early warning system. Previous work has shown that the brain automatically monitors unfolding auditory scenes and rapidly detects new events. Here, we focus on understanding how automatic change detection interfaces with the networks that regulate arousal and attention, measuring pupil diameter (PD) as an indicator of listener arousal and microsaccades (MS) index of attentional sampling. Naive participants were exposed to artificial 'scenes' comprised of multiple concurrent streams of pure tones while their ocular activity was monitored. The scenes were categorized as REG or RND, featuring isochronous (regular) or random temporal structures in the tone streams, respectively. Previous work showed that listeners are sensitive to predictable scene structure and use this information to facilitate change processing. To model changes, a single component was added (change appearance) or removed (change disappearance) from the scenes. Results show that non-attended scene changes elicit pupil dilation and inhibit MS responses, providing evidence for automatic attentional capture and increased arousal. Importantly, change-evoked MS inhibition (MSI) responses were modulated by scene regularity, exhibiting increased MSI in REG scenes, consistent with heightened attentional capture by changes in predictable contexts. Overall, findings shed light on how automatic auditory scene analysis interfaces with attentional/arousal networks.

P5.4 The effect of hearing aid use on phonological processing performance

Ruijing Ning, Emil Holmer & Henrik Danielsson

Linköping, Sweden

Age-related hearing impairment has been shown to affect non-auditory cognitive tasks, such as visual rhyme judgment, lexical decision, and physical matching, which has been attributed to declining phonological representation. While hearing aid use has been found to alleviate the negative effects of hearing impairment on speech perception, its influence on restoring phonological representation remains unclear. Participants with hearing impairment wearing hearing aids (the HA group, n = 214) and those with hearing impairment not wearing hearing aids (the NHA group, n = 71) within the same age range were examined on three tasks involving varying degrees of phonological processing: rhyme judgment, lexical decision, and physical matching. The preliminary findings indicated that the NHA group generally performed worse on these tasks compared to the HA group, particularly on tasks requiring explicit phonological processing. These results suggest that hearing aid use helps in the restoration of phonological processing abilities. Furthermore, the study explored the interaction between hearing aid use and working memory capacity in influencing phonological processing. This research sheds light on the significance of hearing aid use beyond auditory tasks and has implications for the plasticity of the aging brain.

P5.5 The interaction between energetic masking and attentional control during divided-attention listening

Sarah Knight, Georgie Maher & Sven Mattys

University of York (York, UK)

Understanding how listeners track two talkers simultaneously (i.e., divided-attention listening) requires modelling the interaction between bottom-up (acoustic) and top-down (cognitive) factors. The former primarily concern energetic masking (EM) – spectrotemporal interference at the auditory periphery. The latter involve control of auditory attention. In particular, auditory attention may need to be directed to different spatial locations during divided-attention listening. The impact of spatial separation between talkers in this situation is unclear: it is likely to create acoustic benefits (release from EM) but also to increase cognitive costs (increased demands on spatial attention control). In a series of studies, we manipulated: 1) degree of EM between two talkers – high (voices unfiltered) vs. low (voices in non-overlapping frequency bands); 2) spatial separation between talkers – collocated (diotic) to maximally separated (dichotic). When EM was high, transcription performance improved monotonically from collocated to dichotic, indicating spatial release from EM. When EM was minimal, the benefit of separation disappeared and transcription performance actually worsened in the dichotic condition. Additionally, individual differences in working memory best predicted performance in low-EM conditions. These results suggest that acoustic processes dominate during divided-attention listening but that cognitive challenges and contributions can be observed when EM is reduced.

P5.6 Effects of Sleep, Age, and Hearing on Ease and Effectiveness of Communication, and the role of Cognitive Resources

Stephanie Loukieh, Antje Heinrich, Josef Schlittenlacher & Karolina Kluk-de Kort

University of Manchester, UK

Introduction: Older adults face various multifaceted challenges including hearing impairment, cognitive decline, and sleep disruption; and when occurring together, unexpected consequences may arise. This study investigates the effect of factors that may deplete cognitive resources on ease and effectiveness of communication.

Methods: Three experiments will measure the individual and additive effects of sleep quality, age, and hearing on listening effort, fatigue, speech perception, cognitive resources, and overall wellbeing. To annul the effect of one of the depletors, the sample with hearing loss will receive hearing aids, and the outcome variables will be assessed again.

Expected results: Poorer sleep, older age, and hearing loss conditions are expected to negatively affect cognitive resources, especially when co-occurring. A decline in cognitive resources is expected to reflect poorly on listening effort, fatigue, speech perception, and overall wellbeing. Aiding hearing is expected to free up cognitive resources, and reflect positively on the other outcome variables, and indirectly improve sleep.

Implications: If additive effects of multiple depletors are observed and associated with poor outcomes, a preliminary model of cognitive resource depletion will be explored and expanded with other potential depletors.

Keywords: communication, sleep, hearing impairment, listening effort, cognitive resources

P5.7 Understanding listening difficulties in children: Investigating in typically-developing children the relationship between hearing, speech, language, and cognition as a first step to diagnosing APD

Xuehan Zhou[1], Harvey Dillon[1], Alisha Gudkar[1], Kelly Burgoyne[1], Dani Tomlin[2] & Antje Heinrich[1]

[1]The University of Manchester; [2]The University of Melbourne

A range of deficits can cause children difficulty when understanding speech in challenging listening environments, such as noisy classrooms. Children with listening difficulties are at risk of having poor long-term academic outcomes and social skills (Barry et al., 2015), especially when clinicians cannot detect or remediate their specific deficits. Deficits in auditory, speech, language, or cognition abilities may present in a similar manner. Currently, it is difficult to determine the cause of these difficulties. A systematic approach to differentiate between these causes in individual children has been devised (Dillon & Cameron, 2021).

Children aged 6-12 years, enrolled in mainstream primary schools, go through a tri-level test battery; a combination of top-level speech perception in noise and reverberation ability, mid-level phoneme identification ability, and low-level acoustic resolution ability is applied (Dillon & Cameron, 2021). In conjunction with language and cognitive test scores, the combined approach allows for differentiation of the cause of the observed listening deficit.

Speech-sound identification ability in noise and reverberation, non-speech auditory processing abilities, language abilities and cognitive abilities will be used to predict the understanding of sentences in noise and reverberationg for APD.

P5.8 The Irrelevant Speech Effect in Tonal Languages for Mandarin and English native speakers

Yupeng Wu[1], Karolina Kluk-de kort[1], Rebecca Millman[1], Josef Schlittenlacher[2]

[1]University of Manchester; [2]University College London

Background sounds have a disruptive effect on people's ability to memorize information, even if the target is presented visually. Because speech has been found to be the most disruptive sound, this phenomenon was termed the Irrelevant Speech Effect (ISE). Research has focused on various aspects of the ISE e.g., acoustic and language factors, but there is little literature that investigates the ISE for tonal languages. We hypothesize that, at least for native speakers of a tonal language, the ISE is bigger for tonal languages due to the additional feature of pitch changes.

In the present study, we ask native Mandarin and native English speakers to memorize the order of a sequence of nine digits that are presented visually while listening to background sounds. These sounds were either English, Mandarin, or Cantonese speech, or a continuous pink noise. The study is pre-registered at osf.io/63gha. Interim results of 19 native Mandarin speakers confirm that the least disruptive sound is the continuous pink noise, with 8.4 of the 9 digits recalled correctly, on average. The participants recalled 7.0 digits for English background speech, 7.0 for Cantonese and 6.7 for Mandarin. These results show the expected tendency of Mandarin speech to be more disruptive than English. However, Cantonese seems to be less disruptive than Mandarin and equally disruptive as English. Data collection of this study is ongoing, we expect to report 80 participants at the time of the conference.

P5.9 Measuring the predictive ability of various aspects of working memory to speech in noise perception

Antje Heinrich

University of Manchester, UK

In addition to the ability to hear, accurate speech-in-noise perception also requires contributions from cognitive abilities. The cognitive ability that is examined most often in the context of speech-in-noise (SiN) perception is working memory (WM). According to Baddeley and Hitch (1974) two components are critical for WM: a storage component and a manipulation component. WM tasks differ in how they combine these two components. One WM task commonly used in the context of speech perception is the reading span task (RST). Why the RST works so well to predict speech-in-noise perception performance is not well understood. It is also not well understood whether it is the storage or the manipulation component that has the highest predictive value, or whether both components combine to maximise the task's predictive value. Finally, it remains to be understood whether the predictive ability of the RSTs' two components differs either for different groups of listeners or for different speech-in-noise tasks.

We assessed performance of the storage and manipulation components of an RST and related them to the perceptual accuracy of several SiN tasks in two groups of normal-hearing young listeners: English native speakers and English non-native speakers. This is the first experiment in a programme aimed to fully understand the predictive ability of the reading span task for various groups off listeners and tasks and to explore the use of the RST for audiological practice.

P5.10 Frequency Discrimination and Music Enjoyment in Adult Cochlear Implant Users

Cynthia Lam, Nicholas Haywood, Brian C. J. Moore, Ben Williges & Deborah A. Vickers

University of Cambridge

The effectiveness of CIs varies depending on a person's hearing experience. Despite advancements in speech perception, CI users with little to no residual hearing may struggle with perceiving music, partly due to the limited access to low-frequency information. Lam et al. (2022) found that CI users were better in discriminating chords when the second note was lowered than when the third note was raised. This effect could reflect the use of within-channel temporal interactions between tones in the form of beats, which are more salient when the fundamental frequencies of the constituent tones are close together. As CIs primarily provide information about temporal envelope fluctuations in different frequency regions, the detection and discrimination of beats could play a significant role in speech and music perception. This ongoing study aims to examine how CI users discriminate between tones for which the frequency components may fall into one channel (giving rise to beats) or multiple channels (reducing the salience of beats), and to assess the relationship between using beat cues and speech/music perception. We aim to test 24 CI users, by conducting a questionnaire, undertaking frequency discrimination tasks and measuring speech and melody perception. Preliminary results suggest that some CI users are able to use beat cues to discriminate frequency, but not necessarily in all channels and frequency ranges. Overall, this study enhances our understanding of cues that CI users use.

P5.11 Using measures of amplitude modulation processing to understand function of electrically stimulated auditory pathways

Deborah Vickers[1], Brian Moore[1], Ben Williges[1], Patrick Boyle[2], Marina Salorio-Corbetto[1], Josef Schlittenlacher[3] & Nicholas R. Haywood[1]

[1]University of Cambridge; [2]Advanced Bionics GMBH, Hannover, Germany; [3]University College London, London

For cochlear implant (CI) listeners speech information transmission is reliant upon the ability to process the amplitude-modulated (AM) envelope of speech sounds independently in different channels. This can be hindered for many reasons, not least due to, spread of electrical current or neural survival.

We employed a psychoacoustic task to explore AM processing. We recruited normal hearing adults and adult CI listeners (Nucleus and Advanced Bionics). Acoustic sinusoids of two different rates (for example, 13 versus 40 Hz) were discriminated in a three-interval two-alternative forced choice task, where the modulation depth was adjusted adaptively to derive an AM discrimination threshold. Testing was conducted with and without speech envelope interferers on neighbouring channels. Stimuli were delivered through headphones. All front-end noise reduction features were de-activated. We explored AM processing across the frequency range.

Initial findings suggest that AM discrimination was poorer in the presence of interferers for both normal hearing. There is variability across CI listeners and within the dataset of individual listeners. We interpret the measure as indicators of neural function when interferers are absent and indicative of channel interaction when interferers present. We will calculate the normative ranges for the measures. Ongoing work will compare findings to objective measures of viability of the electrode-neurone interface and speech in noise measures

P5.12 Design of a user interface for state-of-the-art psychometric function estimation

Emilie d'Olne[1], Benjamin Levett[2], Jason D. Warren[2] & Patrick A. Naylor[1]

[1]Imperial College London; [2]University College London

The topic of psychometric function (PF) evaluation is an important issue in psychoacoustics and many algorithms have been developed to accurately estimate its speech-reception threshold (SRT) and slope. However, these methods are rarely exploited by practitioners for they typically require coding expertise to be used in trials. In this work we have created a General User Interface (GUI) for PF evaluation which implements the state-of-the-art adaptive Gaussian method in [1] in an easy-to-use manner. The GUI is designed to support several estimation methods and trial types, and it will be used in an upcoming trial to evaluate the effect of reverberation on people affected by dementia syndromes.

[1] Clement S. J. Doire, Mike Brookes, Patrick A. Naylor; Robust and efficient Bayesian adaptive psychometric function estimation. J Acoust Soc Am 1 April 2017; 141 (4): 2501–2512.

P5.13 Temporal Pitch Perception for Selective Apical Stimulation in Cochlear Implant Recipients

Evelien De Groote[1], John M. Deeks[1], Robert P. Carlyon[1] & Olivier Macherey[2]

[1]Cambridge Hearing Group, MRC Cognition and Brain Sciences Unit, University of Cambridge (Cambridge, United Kingdom); [2]Aix-Marseille Univ., CNRS, Centrale Marseille, LMA, Marseille, France

Cochlear implant (CI) recipients typically show poor temporal pitch perception, as revealed by a low upper limit measured on a single electrode. However, both human and animal data show that temporal pitch processing can be improved through selective stimulation of the cochlear apex. The experiments described here investigate rate-pitch ranking in 3 stimulation conditions that one would reasonably expect to affect either the place of stimulation or the spread of excitation, and examine its relationship with the polarity effect (PE), an index of local neural health. Rate-pitch ranking data from 8 Med-El users for (1) single-electrode apical stimulation was compared to (2) single-electrode stimulation of a medial electrode and (3) multi-electrode apical stimulation. We did not find a significant difference regarding the upper limit or the slope of the pitch-rank function between single-electrode apical stimulation on the one hand and single-electrode medial or multi-electrode apical stimulation on the other hand. These findings did not depend on differences regarding the PE measured at apical versus medial stimulation sites. We did replicate the positive correlation between the normalised PE and average threshold, which has previously been taken as evidence that the PE can be used to estimate the survival of auditory-nerve peripheral processes. To conclude, we found no evidence for superior temporal processing through selective apical stimulation in this group of CI recipients.

P5.14 Factors affecting stream segregation in cochlear implant users

Nicholas Haywood, Ben Williges, Marina Salorio-Corbetto & Deborah Vickers

University of Cambridge, UK

Aspects of stream segregation in cochlear implant (CI) users remain poorly understood. In normal hearing (NH), segregation increases as the frequency separation (ΔF) between alternating tones is increased and/or the inter-stimulus interval (ISI) between tones is decreased. However, while stream segregation in CI listeners appears to be influenced by ΔF, ISI has not been found to affect segregation judgements.

In this on-going research, we asked CI listeners to report perceived segregation in stimuli where both ΔF (i.e., targeted electrode) and ISI were varied – the range of ISIs tested extend beyond those tested previously. In the preliminary dataset, all listeners show an effect of ΔF, and some do show ISI effects.

A second task required listeners to detect a temporal delay imposed on a single tone. Stimuli were arranged so that any obligatory stream segregation should impair performance.

Preliminary results are varied, all listeners showed an influence of ΔF indicative of stream segregation, but the way a segregation-promoting precursor sequence affected performance varied across listeners.

We will compare these measures of stream segregation to other aspects of auditory performance – such as speech in noise.

P5.15 Acoustic cues for biological sound-source perception by human listeners.

Nicole Miller-Viacava[1], Anne C. Axel[2], Regis Ferriere[3,4,5], Nicholas R. Friedman[6], François-Michel Le Tourneau[3], Diego Llusia[7,8,9], Timothy C. Mullet[10], Yvonne F. Phillips[11], Jacob Willie[12,13], Jérôme Sueur[14] & Christian Lorenzi[1]

[1]Laboratoire des systèmes perceptifs, École normale supérieure, PSL University. Paris, France; [2]Department of Biological Sciences, Marshall University, Huntington, WV 25755, USA; [3]International Research Laboratory for Interdisciplinary Global Environmental Studies (iGLOBES), CNRS, ENS-PSL University, University of Arizona, Tucson, AZ, 85721, USA; [4]Institut de Biologie de l'Ecole Normale Supérieure (IBENS), Université Paris Sciences et Lettres, CNRS, INSERM, Paris, 75005, France; [5]Department of Ecology & Evolutionary Biology, University of Arizona, Tucson, AZ, 85721, USA; [6]Environmental Informatics Section,Okinawa Institute of Science and Technology Graduate University, Onna, Japan; [7]Terrestrial Ecology Group (TEG), Department of Ecology, Universidad Autónoma de Madrid (UAM), Madrid, Spain; [8]Center for Research on Biodiversity and Global Change (CIBC-UAM), Universidad Autónoma de Madrid, Madrid, Spain; [9]Programa de Pós-Graduação em Biodiversidade Animal (PPGBAN), Laboratório de Herpetologia e Comportamento Animal (LHCAN), Instituto de Ciências Biológicas, Universidade Federal de Goiás (UFG), Goiânia, Brazil; [10]Kenai Fjords National Park, U.S. National Park Service, Seward, AK, United States; [11]Science and Engineering Faculty, Queensland University of Technology, Queensland, Australia; [12]KMDA, Centre for Research and Conservation, Royal Zoological Society of Antwerp, Koningin Astridplein 20-26, B-2018 Antwerp, Belgium; [13]Terrestrial Ecology Unit (TEREC), Department of Biology, Ghent University (UGent), K.L. Ledeganckstraat 35, 9000 Ghent, Belgium; [14]Institut de Systématique, Évolution, Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, 57 rue Cuvier, 75005 Paris, France

The aim of this study was to characterise the acoustic cues used by human listeners when perceiving sounds produced by biological sources (biophony) in congruent settings. Previous studies have identified spectro-temporal features that may play a key role in this perceptual process. However, their stimuli had poor ecological validity due to, amongst others, an over-representation of mammals and lack of consideration of sound-propagation effects in natural environments. To address these issues, a database was constructed using natural soundscapes from nine different habitats with minimal human activity, for two seasons and four moments of the day. One-sec long samples were used to assess the capacity of human participants to categorize each sound source as biological or geophysical. Sounds in each category were analyzed by computing their modulation power spectrum and deriving several modulation statistics from it. Preliminary results show that biophony categorization was mainly driven by presence of low spectral modulation frequencies (<4 cyc/kHz) for a large range of temporal rates. In agreement with previous findings, the contribution of slow temporal rates (<10 Hz) expanded up to higher spectral modulation frequencies (<8 cyc/kHz), a potential indicative of the role of harmonic

sounds with rather high fundamental frequencies. Additional analyses and psychophysical experiments are in progress to unveil further details of which cues are exclusive of biophony perception.

Session 6: Multi-Modal Hearing and Hearing-Assistive Technologies

Chair: Amir Hussain (Edinburgh Napier)

Oral presentations

O6.1 Developing a haptic hearing-assistive device to improve speech perception in hearing-impaired listeners

Samuel Perry & Mark Fletcher

The University of Southampton, UK

In-lab research has shown that hearing can be augmented using haptic stimulation presented on the wrists to improve speech perception. This method could provide benefits for hearing-impaired listeners who struggle in areas such as speech-in-noise performance, pitch discrimination and sound localisation. For example, this low-cost technique may particularly benefit cochlear implant (CI) users and listeners who are unable to access CI technologies, such as those in developing countries. We are developing a wrist-worn device to provide these benefits in real-world environments.

To assess the efficacy of the device in a real-world setting, we aim to accurately replicate the stimulation methods used in previous in-lab studies, using substantially more limited hardware. To achieve this, the wristband design and motor performance should be optimised, with consideration for the physical and psychophysical limits of the device. In this work we used non-linear system identification techniques (e.g., Hammerstein models) to drive a control system, which aims to correct for distortions in the temporal and frequency response of the device. The faithful production of tactile stimulus is likely to be critical to reproducing previous improvements observed in lab. Therefore, the presented work may contribute to the realisation of a high-fidelity haptic wristband that provides real-world improvements in speech, sound localisation and music perception for hearing-impaired listeners.

O6.2 Searching for individual differences in audiovisual integration of speech in noise

Chris Sumner[1], Samuel Smith[2], Jens Roeser[1], Thom Baguley[1] & Paula Stacey[1]

[1]Nottingham Trent University; [2]Massachusetts Eye and Ear Infirmary, Boston, USA

Speech comprehension is often aided by watching a talker’s face. This is of particular value to those with hearing impairment. It is clear that there are individual differences in the ability to understand auditory speech in noise, and differences in the ability to understand visual speech (“lip-reading”) although most people are poor at this. It remains an open question whether people differ in the manner in which they integrate auditory and visual cues. We do not know what the “integration function” is, and whether it is the same for everyone. To address this question, we have applied a new method for quantifying multi-sensory integration, based on signal detection theory (SDT), to a large dataset of audio-visual speech perception performance. Participants (>150) vary in (self-reported): age, hearing-loss, English language experience, language-specific impairments and neurodiversity. Audiovisual performance was for the most part accurately predicted by unisensory performance. In contrast, for both auditory-only and visual-only speech perception there were striking differences in individual performance. There was also some evidence of some small systematic differences in auditory performance across the demographic groups, and potentially in the integration function. This suggests that the "audiovisual integration function" for speech is relatively consistent across a diverse population, with individual differences being attributable to differences in unisensory perception.

O6.3 How does vision benefit speech perception in noise?

Lida Alampounti, Hannah Cooper & Jennifer Bizley

University College London (UCL), London, United Kingdom

Lipreading benefits listening in noise. Audio-visual temporal coherence can also help listeners segregate competing sounds.

This study's aim was to assess whether the benefit of seeing a talker's face is fully explained through lipreading, or whether visualising mouth movements also provides temporal coherence cues that may help listeners segregate competing sources.

We used a speech-in-noise task involving the identification of two target words in a sentence in the presence of two masker speakers. We measured speech discrimination thresholds while presenting the sound mixture accompanied by a naturalistic video of either the target talker or one of the masker talkers (naturalistic). We separately measured participants' ability to use vision for streaming by presenting the video of either the target talker or a masker but freezing the talker's mouth for the duration of the target words (interrupted). Finally, we measured performance with audio and a static image of either the target talker or a masker (static image). In addition to speech-in-noise testing, we measured participants' hearing thresholds and ability to lipread.

Results from 125 participants (aged 18-85, 88 hearing, 37 with hearing loss) suggest that both naturalistic and interrupted offer an advantage over the static image in hearing listeners. Analysis is ongoing, but preliminary results suggest that increasing age and/or hearing loss diminish the advantage offered in the interrupted condition.

O6.4 End-to-End demonstration of an Audio-Visual Speech Enhancement Model in an Embedded Edge AI System

Islam Zakaria Nasr, Kia Dashtipour, Mandar Gogate & Amir Hussain

Centre of AI and Robotics, School of Computing, Engineering and Built Environment, Edinburgh Napier University (Edinburgh, Scotland, UK)

Numerous audio-only (AO) and audio-visual (AV) speech denoising models, especially those designed for hearable devices, have often overlooked the deployment constraints imposed by limited-resource devices, such as memory and processing units. As a result, our research objectives encompass the following key aspects:

1. Constructing an end-to-end demo showcasing cutting-edge speech enhancement techniques tailored for embedded devices, including seamless integration of microphones to ensure optimal performance.

2. Sharing tools encompassing the entire development cycle, including data collection, model training (AutoML), and model optimization using technologies like TensorFlow Lite Micro (TFLu).

3. Propagating best practices for effective preprocessing and postprocessing of AO and AV data, with a strong emphasis on achieving software design harmony for enhanced efficiency and usability.

We present a complete life cycle production of AO and AV speech enhancement, comprising eight stages: Data Collection, Preprocessing, Model Design, Model Training, Evaluation & Optimization, Model Conversion, Model Deployment, and Making Inferences. Each stage provides valuable insights into the encountered challenges and the specific tools employed, offering a holistic view of the demo. This demonstration serves as a pivotal milestone, driving future advancements in audio-visual speech enhancement, particularly for hearing aids and similar applications.

Poster presentations

P6.1 The next way forward for Audiology – Hearing Assistive Devices that remodel the brain.

James Mander

Ewing Foundation. Based within The University of Manchester. UK.

Vibrotactile (VT) devices help listeners with moderate to severe hearing loss. These devices deliver vibrations, representing sound and speech to the surface of the skin. VT devices deliver tactile stimulation in real-time thereby using the brain's ability for plasticity to learn new associations to interpret these sensations (Perrotta et al 2021). An investigation by Fletcher et al (2018) demonstrated a 10.8% improvement in words. In principle VT devices should benefit all age groups. We propose a study using these devices on range of adult users groups with hearing loss who wear hearing devices ranging from HAs to CIs to BCHDs. The study will examine the following questions: 1.0) Does wearing a VT device in addition to a hearing aids (CIs or BCHD) improve listeners' speech recognition performance in quiet and in noise? 1.1) Does the device improve environmental sound awareness and identification? 2) How long is the acclimatisation period for such a VT until peak performance is achieved? 3) Are improvements in speech perception also reflected in increase Quality of Life in everyday situations? The study would analysis the results especially from different user groups with HAs, CIs, BCHDs. The outcomes would help to determine the utility of such devices for deaf users; provide information on cost effectiveness; and provide information on which groups of listeners, in terms of hearing instrument usage and age, might benefit the most.

P6.2 Applications of automatic speech recognition and text-to-speech models to detect hearing loss: a scoping review

Mohsen Fatehifar[1], Josef Schlittenlacher[2], David Wong[3], Timothy Cootes[1] & Kevin J Munro[1]

[1]*The University of Manchester (Manchester, UK);* [2]*UCL;* [3]*Leeds Institute of Health Sciences, University of Leeds*

AIMS: 1) Scope recent studies using Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) techniques for hearing assessment. 2) Categorise studies based on their application.

METHODS: The research protocol was pre-registered (inplasy.com/inplasy-2023-1-0029). A search of ten databases (terms included ASR, TTS, machine learning, hearing loss and hearing assessment) was conducted January 2023 with no limitation on publishing date. The search revealed 1578 matches, 1537 of which were excluded based on title and abstract, and a further 22 excluded after reviewing the full text. A total of 19 studies were included.

RESULTS: Four studies used TTS to replace recorded stimuli and none of the synthetic stimuli impacted negatively on accuracy. Six studies used ASR to capture the participants' verbal responses. These ASR systems provided reliable Speech Reception Thresholds. Three studies investigated using ASR to optimize hearing aid configurations. Results show improved ASR accuracy and comfort level for human participants compared to clinical procedures. Finally, seven studies simulated hearing tests. Results showed strong correlations between simulated and empirical data.

CONCLUSIONS: The findings (scoping review manuscript in preparation) show that ASR and TTS systems have the potential to make hearing tests more accessible and reduce the effort needed to create and conduct hearing tests. We aim to create a hearing test with ASR and TTS that is closer to natural conversations.

P6.3 A complete computer model of the auditory periphery

Ondrej Tichacek[1], Pavel Mistrík[2] & Pavel Jungwirth[1]

[1]*Institute of Organic Chemistry and Biochemistry, Prague, Czech Republic;* [2]*Med-EL, Innsbruck, Austria*

We present a computer model of the outer and middle ears, cochlear mechanics, ionic currents within the organ of Corti, the ribbon synapse of the inner hair cells, and the axon of the afferent auditory nerve. The model translates into a large set of differential equations that we solve numerically in the time domain, simulating the response to arbitrary sounds.

The present model is well physiologically justified, built bottom-up using laws of physics, and based on up-to-date experimental data. The meticulous design of the model facilitates the exploration of specific aspects of hearing mechanisms and their deficiencies, encompassing genetic, molecular, and microscopic factors. Additionally, the model enables the examination of microscopic influences within the broader context of their impact on overall hearing. This allows for correlations with neural recordings and non-invasive psychoacoustic techniques.

The model has the potential to serve as a virtual diagnostic tool, enabling the identification of pathological mechanisms associated with various types of hearing impairments. Specifically, the model can simulate the disorder originating from different parts of the auditory pathway, such as abnormalities in mechanoelectric transduction, the electromechanic response of OHCs, neurotransmitter release in IHCs, and spike generation in the auditory nerve. This allows the model to assist in the diagnosis of hearing loss and potentially differentiate between its primary causes.

P6.4 Transport of therapeutic agents along the cochlea by acoustic steady streaming - 3D FEM study

Xinyu Zhou & Torsten Marquardt

UCL Ear Institute

With the recent development of inner ear therapies, pharmaceutical agents will soon become clinically available. But the application of the drug to the inner ear still faces challenges. The blood-labyrinth barrier makes systemic administration for many compounds impossible, and trans-tympanic application on or through the round window is the only option. Once inside the cochlea, they still need to be transported towards the apex. Acoustic streaming is an non-linear acoustic effect that has the potential to facilitate the homogeneous distribution of therapeutic agents once inside the scala tympani. The steady flow driven by acoustic stimulation is quantified by the non-linear term of the Navier-Stoke equation. Streaming eddies been observed already decades ago in physical cochlear models, and more recently simulated in simple two-dimensional finite-element models (FEM). Using the FEM package COMSOL Multiphysics and applying the perturbation technique, which hugely reduces computational cost, we were able to simulate streaming in a three-dimensional model. We will discuss the difference between the 2D and 3D streaming results, demonstrate how multiple tones can elongate the streaming eddy along the length of the cochlea and how particle tracing can simulate the path of the compounds.

P6.5 Speech envelope coding in computer-modeled auditory nerve fibers

Mengchao Zhang[1] & John Culling[2]

[1]Aston University (Birmingham, UK); [2]Cardiff University

Auditory nerve fibers (ANFs) with different spontaneous rates (SRs) show limited but complementary dynamic ranges which are thought to handle the wide dynamic ranges of human hearing. Recent physiological studies demonstrate adaptive properties of ANFs wherein the rate-level functions of the ANFs shift according to the overall sound level statistics so as to avoid firing rate saturation at high sound intensities. These findings prompt the question whether dynamic range adaptation alone is sufficient for ANFs to encode speech in noise at various sound levels. The current study uses Meddis' Model of the Auditory Periphery and a neural decoder (Meddis et al., 2013; Grange et al., 2022), which can reproduce this dynamic range adaptation. We examined whether ANFs with different SRs are required to encode speech. 30,000 model ANFs (low, medium, and high-SR fibers proportioned 15%, 25%, and 60%) were set up across 30 best frequencies from 56 to 8000 Hz. Responses to speech in modulated noise were compared when fibers from one or two SR groups were removed. When the low-SR fibers are removed (on their own or along with the medium- or high-SR fibers), the speech envelope was still substantially distorted. Despite the dynamic range adaptation, removing medium-SR fibers, high-SR fibers, or both, impacted the speech representation to a much lesser extent.

P6.6 Towards development of a multi-metric audio visual speech in noise test with use of virtual reality

Adeel Hussain[1], Usman Anwar[2], Mandar Gogate[1], Kia Dashtipor[1], Adele Goman[1], Tughrul Arslan[2], Mathini Sellathurai[3], Michael Ackeroyd[4] & Amir Hussain[1]

[1]Edinburgh Napier University, Scotland; [2]Edinburgh University; [3]Herriot-Watt University; [4]University of Nottingham

Intelligibility measures are widely used metrics to measure speech perception in clinical and research settings. One notable metric that is not encompassed by these measures is listening effort (LE) which could potentially provide additional benefit in both clinics and research. The utilisation of pupillary response recordings has been widely used in various research studies to evaluate the level of LE during a speech-in-noise task. Specialised eye-tracking equipment can be used to obtain these measurements. Although eye-tracking equipment is reliable it can be costly to incorporate into all labs and clinics. With the development of technology, virtual reality (VR) has emerged as a next-generation technology which has made its break into the health industry and can be used to further enhance hearing technology. With the use of a VR headset, we aim to develop an automated audio-visual speech in noise test. The multiple sensors integrated within the VR headset that will objectively measure listening effort whilst the individual will provide subjective speech intelligibility scores. Overall, this multi-metric method may provide an important basis for developing a standardised AV speech test that includes pupillometry in both clinic and research utilising VR headset.

P6.7 Wearing multi-modal hearing aids

Dorothy Hardy[1], Michael A Akeroyd[2] & Amir Hussain[1]

[1]Edinburgh Napier University; [2]University of Nottingham

Hearing aids with audio-visual inputs could select and single out speech in noisy environments through use of cameras. Though tried repeatedly in the past, modern technology could revolutionise them. These multi-modal hearing aids could make it easier to hear one voice when several people are talking. The cameras and associated hardware will need to be worn or carried, which adds complexity to existing hearing aid designs. This research seeks to find how hearing aid users might prefer to wear sensors and associated hardware in conjunction with hearing aids. Our aim is to find out if the improvements in speech intelligibility that the additional equipment gives are sufficiently beneficial for users to consider it worthwhile to wear the additional sensors. Software and hardware prototypes were shown to hearing aid users and audiologists to elicit feedback about the benefits and difficulties of using this proposed new multi-modal hearing technology. Initial discussion focused on privacy aspects of use of cameras with hearing aids, indicating that the new technology would be preferred for use in noisy environments where cameras are already used, such as busy, public spaces. Future multi-modal hearing aids will need to be trusted to preserve privacy at the same time as being sufficiently useful in selecting chosen speech in noise; comfortable; and discreet, if they are to compete successfully with existing designs.

P6.8 A VR enabled multi-sensory dataset for speech enhancement evaluation

Jasper Kirton-Wingate, Adeel Hussain, Chong Tang, Usman Anwar, Mandar Gogate, Kia Dashtipour, Ankit Gupta, Mohsin Raza, Mathini Sellathurai, Tughrul Arslan, Tharm Ratnarajah, Qammer Abbasi, Muhammad Imran, Peter Bell, Michael Akeroyd, Adeel Ahsan & Amir Hussain

Edinburgh Napier University

Amongst the various different methods for Speech Enhancement (SE), Multi-modal SE offers superior performance in speech intelligibility and quality, especially in low SNR scenarios. For the design of successful Hearing Aids (HA) in complex listening environments, it is also important to allow the HA user to experience the myriad of other physical interfaces we use for hearing, such as visual cues and head orientation. Importantly, it is also of concern to evaluate the mitigation of Cognitive Load that may occur with successful SE and associated aiding paradigms. Numerous datasets exist to train, test and evaluate audio-visual SE, however VR enabled datasets are limited, especially amongst real recordings in life-like noisy conditions. Additionally, datasets with RF signals for privacy preserving facial features and context estimation in multi-modal SE are lacking. As a result, the proposed dataset allows the user to engage in a VR experience of two-target speaker listening, whilst enabling the evaluation of multi-modal SE models within a VR HA context. The sentences chosen to be spoken were taken from British IEEE Harvard sentences. The sentences are employed in a turn-taking scenario and a more challenging sentence overlap scenario with simultaneous speech. The target speaker can be chosen by the head orientation angle, or by other measures that predict attention, such as eye tracking. Other sensors can also be employed to assess the CL effects of aiding through VR headsets.

P6.9 Radio Frequency based Cognitive Load Detection for Multimodal Hearing-aids

Usman Anwar[1], Tughrul Arslan[1], Adeel Hussain[2], Mandar Gogate[2], Kia Dashtipour[2], Adele Goman[2], Mathini Sellathurai[3], Michael Ackeroyd[4] & Amir Hussain[2]

[1]University of Edinburgh, Edinburgh, United Kingdom; [2]Edinburgh Napier University; [3]Heriot-Watt University; [4]University of Nottingham

One-third of people over sixty-five years of age are affected by disabling hearing loss. Hearing impairment can lead to cognitive impairment and neurodegeneration in most cases. Listening efforts to comprehend speech result in higher utilization of cognitive resources for hearing-impaired individuals. Cognitive load (CL) represents the strain on cognitive resources such as working memory, brain processing, and visual and verbal information processing units. Cognitive strain eventually leads to disruption of brain blood flow to secondary areas of cognition, affecting cerebral metabolism. Cognitive load can be measured precisely through the detection of blood flow variation in the circle of Willis, white, and gray matter areas of the brain. A proof-of-concept model for real-time cognitive load detection is presented in this work. An innovative radio frequency (RF)-based CL sensing approach is proposed, which can successfully detect high and low cognitive load states based on pathophysiological measurements. The prospect of integrating it with multimodal hearing-aid devices is explored. These multimodal hearing-aid devices will be able to assess the listening effort through cognitive load estimation. Speech enhancement algorithms will be integrated with the multimodal RF sensors to enhance the listening experience of hearing-aid users by translating cognitive load measurements to a listening effort metric.

Session 7: Effects of Noise Exposure on Hearing and Therapeutic Training

Chair: Chris Plack (University of Manchester)

Oral presentations

O7.1 Hearing Assessment of Music Conservatoire Students

Stephen Dance[1], Ruben Vazquez Amos[1] & Georgia Zepidou[2]

[1]London South Bank University, London, UK; [2]AECOM

Since the introduction of the UK's Control of Noise at Work Regulations research has been undertaken in collaboration with the Royal Academy of Music investigating the hearing acuity of more than 5000 students between 2007 and 2023. A standard audiometric screening method was employed for both entry and exit testing of undergraduate and postgraduate students. Since 2021 600 musicians have had their audiometric screening based test supplemented with otoacoustic emission based hearing assessment using Hearing Coach firmware/software to identify damage in the outer hair cells. Results showed that otoacoustic emissions were able to identify, at an early stage, hearing damage in more individuals compared to audiometry. The otoacoustic emission test method was found to be quicker, more convenient and offered greater reproducibility. This provides reassurance that otoacoustic emissions is an excellent supplementary tool for assessing the hearing health of classical music students.

O7.2 Blood prestin levels following music exposure that induces temporary threshold shifts

Eleftheria Iliadou[1,2], Christopher J. Plack[3,4], Konstantinos Pastiadis[1], Athanasios Bibas[1]

[1]National and Kapodistrian University of Athens, Athens, Greece; [2]Ear Institute, University College London, London, UK; [3]Manchester Centre for Audiology and Deafness, University of Manchester, Manchester, UK; [4]Department of Psychology, Lancaster University, Lancaster, UK

Objectives: To determine if blood prestin levels change after exposure to loud music, and if the change is associated with temporary threshold shift (TTS) or a decrease in distortion product otoacoustic emission (DPOAE) amplitude. Methods: Fourteen adults [nine women; median age=31, IQR=6.75] with normal hearing were exposed to music (15’, 100 dBA). Pure tone audiometry, DPOAE amplitude, and blood prestin levels were measured before and after exposure. Results: There was significant average TTS at 4’ post-exposure. Significant DPOAE shifts were observed at 4 kHz, and 6 kHz. Mean baseline prestin (mean=140.00pg/ml, 95% CI: 125.92, 154.07) progressively increased following music exposure, reaching maximum at 2 hours (mean=158.29pg/ml, 95% CI: 130.42, 186.66), and returned to pre-exposure levels at 1 week (mean=139.18pg/ml, 95% CI: 114.69, 163.68). After correction for multiple comparisons, mean prestin levels showed no significant increase from baseline at any timepoint. Although no correlation between maximum blood prestin change and average TTS or DPOAE amplitude shift was found, in exploratory analysis, TTS at 6 kHz (where maximum TTS occurred) decreased significantly as baseline prestin levels increased. Conclusions: Blood prestin levels may change after exposure to loud music, although statistical significance was not reached in our relatively small sample after correction. The role of blood prestin as marker for temporary cochlear dysfunction should be further explored.

O7.3 Do auditory brainstem responses reflect cochlear synaptopathy - evidence from frequency-specific responses

Katrin Krumbholz[1], Alexander Hardy[2], Eden Burch[3] & Jessica de Boer[4]

[1]University of Nottingham, Nottingham, UK; [2]Nottingham Trent University; [3]St-Andrews University; [4]UK Health Security Agency

It has been suggested that noise exposure can cause irreversible loss of cochlear afferent synapses (“cochlear synaptopathy”, CS) despite normal audiometric thresholds. Attempts at demonstrating CS in humans have often involved transient-evoked auditory brainstem responses (ABRs). Some have shown reduction in ABR amplitude, particularly of wave I, in subjects at particular risk of CS. It has been argued, however, that these effects may have been confounded by contributions from audiometrically uncontrolled extended high-frequency (EHF) cochlear regions. It has even been suggested that ABRs may be inherently insensitive to the type of synapses (low-spontaneous-rate synapses) thought to be most affected by CS. The current study explores a data set comprising both aggregate ABRs, like those used in previous studies, and frequency-specific ABRs (so-called “derived bands”, or DBs) arising from restricted cochlear regions. Subjects exhibited normal hearing thresholds within the standard audiometric range, but a large range of noise-induced variation in EHF thresholds. The DB amplitudes showed a clear and systematic dependence on sub-clinical variation in relevant local hearing thresholds, whilst, at the same time, also showing evidence of audiometrically hidden damage. Using a phenomenological model of ABR amplitude growth based on level-dependent DB-ABR data, we explore whether the latter reflects CS or not-yet-audiometrically-manifested decline in hair cell function.

O7.4 Perception and Measurement of Audio Quality Attributes in Music by Hearing Aid Users: A Sensory Evaluation Study

Alinka Greasley[1], Scott Bannister[1], Trevor Cox[2], Michael Akeroyd[3], Jon Barker[4], Bruno Fazenda[2], Jennifer Firth[3], Simone Graetzer[2], Gerardo Roa Dabike[2], Rebecca Vos[2] & William Whitmer[3]

[1]University of Leeds; [2]University of Salford; [3]University of Nottingham; [4]University of Sheffield

Hearing aids (HAs), optimized for speech perception, can be problematic for music given its broader and variable spectral and dynamic characteristics. This study aimed to improve understanding of music audio quality for HA users and develop audio quality metrics. Twelve HA users completed an elicitation task, providing up to 3 terms to describe audio quality of 27 music samples. They discussed these terms across 3 focus groups, to reach a consensus on perceptual attributes and their definitions. The elicitation task resulted in 373 perceptual terms with the most used including clear, loud, distorted,

balanced and unclear. In focus group 1, discussions began around terms used 4 or more times (N=89). Participants listened to samples and chose one term that best described each; 45 terms were selected. They then discussed similarities and differences between these, creating loose groupings. In focus group 2, the meaning of the groups was discussed, resulting in 7 perceptual term structures. In focus group 3, participants agreed on labels, definitions, and rating scales, resulting in 7 perceptual attributes of clarity, harshness, distorted, treble, middle and bass strength, and spaciousness. Whilst work has explored audio quality perception in 'normal' hearing listeners, research with hearing-impaired listeners is scarce. This study has explored important dimensions of music audio quality relevant for HA users, developing perceptual attributes of audio quality for future research.

Poster presentations

P7.1 A systematic review of measurements of real-world interior car noise for the Cadenza machine-learning project

Jennifer Firth[1], Scott Bannister[2], Jon Barker[3], Trevor Cox[4], Bruno Fazenda[4], Simone Graetzer[4], Alinka Greasley[2], Gerardo Roa Dabike[4], Rebecca Vos[4], William Whitmer[1] & Michael Akeroyd[1]

[1]University of Nottingham; [2]University of Leeds; [3]University of Sheffield; [4]University of Salford

Interior car noise refers to the general noise generated by the engine transmission, the interaction between road and types, and weather conditions such as turbulent wind. For drivers or passengers with hearing loss, these can create especially challenging listening situations. The Cadenza Project is organising a series of machine learning challenges to advance signal processing of music for listeners with a hearing loss. A key scenario in its first challenge is listening in a car to music in the presence of noise. To create sufficient machine-learnable training materials, we applied a data-driven approach to simulate typical interior car noises. We conducted a systematic review to identify real-world interior car noise recordings, to determine the parameters for the simulations. The following search terms were applied to Web of Science: "car noise, car noise interior, interior noise AND speed OR FFT OR spectr*". Of a total of 126 papers, 6 were retained on the basis that a 1/3 octave frequency spectrum for interior car noise was provided that was suitable for numerical analysis. In all, 13 real-world recordings were used to generate a stimulated interior car noise, to be used as part of the Cadenza machine-learning project.

P7.2 Developing and validating virtual-audio clinical tools for assessing spatial-listening skills for children with bilateral cochlear implants

Bhavisha Parmar, Marina Salorio-Corbetto, Jennifer Bizley, Stuart Rosen, Tim Green, Lorenzo Picinali, Ben Williges & Deborah Vickers

University of Cambridge

Background

Clinical tests for the assessment of spatial listening require multi-speaker arrays rarely available in clinical settings. A virtual-audio version of the Spatial Speech in Noise Test (SSiN) leads to similar performance across spatial locations for loudspeaker arrays with normal-hearing listeners. The aim of this work is to determine whether the virtual-audio versions of the SSiN and the Adaptive Sentence List (ASL) using a spatial release from masking test configuration test yield comparable results than their loudspeaker versions for children with bilateral cochlear implants. Additionally, the efficacy of a centralisation app to identify the degree of balance between the ears was explored together with the findings from the virtual speech tests. The purpose of this work is to validate virtual assessments for use in a clinical trial with virtual reality spatial training games.

Method

A participatory-design approach was used to develop and finalise the virtual-audio implementations of the tests. Ten children and young adults who wear bilateral cochlear implants and ten age-matched normal-hearing participants, will perform each test (SSiN and Spatial ASLs) in each implementation (virtual-audio or loudspeaker). The order of the tests and implementations were counterbalanced across participants. The participants also completed the centralisation task (i-balance app) using narrow-band noise and wide-band stimuli consisting of speech-shaped noise and a non-language specific speech-like stimulus. The interaural level differences for these stimuli were varied by the children using a visual/tactile interface so that the sound was perceived in the midline. Children were asked to show where they located or heard the sound relative to their head by colouring a drawing.

Results

So far, the virtual-audio applications were finalised. Eight children and young people with cochlear implants, and three with normal hearing, have completed the tests. Our outcomes will allow us to determine whether the virtual-audio versions of the tests have potential for clinical use, provide the validation for use in the clinical trial and determine whether the

results from the centralisation task used in the i-balance app are informative in terms of spatial hearing abilities for children with bilateral cochlear implants.

P7.3 Demo: Bistro-in-a-box noisy environment simulator for evaluation of, and training with, hearing technology

Alastair H. Moore

Imperial College (London, UK)

Difficulty in understanding speech in noise is the most common complaint of people with hearing loss. Hearing aid technology is improving all the time, hearables are becoming commonplace and assistive listening devices, such as a remote microphone, can offer substantial signal-to-noise ratio benefits. However, many users, whether they have a diagnosed hearing loss or not, are reluctant to use such technology to improve their ability to communicate. Often what is lacking is the opportunity to try and/or practice using such technology in a low-stress environment. The Bistro-in-a-box is a wireless loudspeaker array which is easy to setup and control, allowing audiologists, hearing coaches, retailers and service providers to demonstrate the benefits of hearing technology to potential users. Equally, it could be used in the home, allowing hearing technology users to practice using their devices and helping them to become familiar with their limitations.