Hanse-Wissenschaftskolleg
Institute for Advanced Study

Hearing (musical) scenes
international symposium on auditory scene analysis

Delmenhorst, March 27-29, 2019

Preliminary program

as of 01.03.2019

Organizer:
Dr. Kai Siedenburg
Carl von Ossietzky Universität Oldenburg

Venue:

Hanse-Wissenschaftskolleg
Institute for Advanced Study
Lehmkuhlenbusch 4
27753 Delmenhorst
Germany
www.h-w-k.de

Program

supported by

Deutsche Forschungsgemeinschaft
German Research Foundation

Carl von Ossietzky Universität Oldenburg

Wednesday, March 27, 2019

15:00 **Optional lab tour,** University of Oldenburg, Wechloy Campus

18:00 **Dinner at the restaurant** 'Zum Drögen Hasen', walking distance from Wechloy Campus (depending on confirmation by host)

20:00 Bus shuttle to the City Hotel in Delmenhorst

Thursday, March 28, 2019

09:00 **Opening remarks**

09:15 **Keynote 1: Brian C.J. Moore.** Effects of hearing loss and hearing aids on the perception of music

10:00 **Session 1: Hearing loss in music and speech**
Chair: Kai Siedenburg
Frank A. Russo. *Using the Adaptive Music Perception (AMP) test to understand the effects of hearing loss and hearing aids on the perception of musical scenes*
Birger Kollmeier. *How many factors do we need to describe and model suprathreshold "distortion"?*

11:00 **Coffee break**

11:30 **Session 2: Advances in hearing devices**
Chair: Marc Réne Schädler
Simon Doclo. *Binaural speech enhancement and cue preservation algorithms*
Niels H. Pontoppidan. *Compensation for impaired auditory scene analysis*
Volker Hohmann *& Giso Grimm. Multi-modal hearing devices and their evaluation*

13:00 **Lunch break**

14:00 **Session 3: Stream formation**
Chair: Emily Coffey
Georg M. Klump*. A comparative view on auditory scene analysis and informational masking*
Jeremy Marozeau*. Auditory streaming cues in cochlear implant listeners*

15:00 Coffee break

15:30 **Session 4: Multistability**
Chair: Steven van de Par
Alexandra Bendixen*. Multistability in auditory scene analysis*
Kai Siedenburg. *Ambiguity of spectral envelope shifts resolves with prior context*
Jackson Graves *& Daniel Pressnitzer. Binding in auditory scene analysis: Insights from behavioral and pupillometric experiments*

17:00 Short break

17:15 **Keynote 2: Josh H. McDermott.** Computational auditory scene analysis as causal inference

18:00 Adjournment

18:30 Bus shuttle to the 'Bremer Ratskeller'

19:15 Dinner at the restaurant 'Bremer Ratskeller'

21:30 Bus shuttle returns from the restaurant to the hotel

Friday, March 29, 2019

09:15 **Keynote 3: Mounya Elhilali.** Attention at the cocktail party: cognitive control of auditory processing

10:00 **Session 5: Attention**
Chair: Kai Siedenburg
Jannis Hildebrandt*. Attentional capture in auditory scenes: salient events shape perception by improving the neural representation of the capturing stream*
Sophie Nolden *et al. Auditory selective attention in noisy environments – differences between children and young adults*

11:00 Coffee break

11:30 **Session 6: Musical scenes I**
Chair: Stefan Weinzierl
Stephen McAdams. *Scene analysis as a psychological foundation for musical orchestration*
Trevor Agus. *A review of sound engineers' advice for manipulating complex soundscapes*
Jason Noble. *Analyzing assimilated musical scenes: sound mass perception and its acoustical correlates*

13:00 Lunch break

14:00 **Session 7: Musical scenes II**
Chair: Henning Schepker
Waldo Nogueira. *Making music more accessible for cochlear implant listeners*
Clemens Wöllner. *Time dilations in music and audiovisual media*

15:00 Coffee break

15:30 **Session 8: Are musicians special?**
Chair: Kirsten Wagener
Gunter Kreutz et al. *Incidence and relative risk of hearing disorders in professional musicians*
Sarah M.K. Madsen. *Speech perception is similar for musicians and non-musicians under wide range of conditions*
Emily Coffey. *Hearing-in-noise perception: why are musicians better?*

17:00 Short break

17:15 **Special guest: Robert J. Zatorre**. Musicians at the cocktail party: neural correlates of top-down and bottom-up segregation mechanisms

18:00 Closing

18:30 - 20:00 Dinner at the HWK

KEYNOTES

Brian C. J. Moore
University of Cambridge

Effects of hearing loss and hearing aids on the perception of music

Sensorineural hearing loss produces a variety of physiological changes in the auditory system that in turn produce a variety of perceptual effects. Damage to the outer hair cells within the cochlea leads to a loss of sensitivity to weak sounds, loudness recruitment (a more rapid than normal growth of loudness with increasing sound level and a consequent reduced dynamic range), and reduced frequency selectivity. Damage to inner hair cells and/or synapses can lead to degeneration of neurons in the auditory nerve and hence to a reduced flow of information to the brain, even when audiometric thresholds remain normal. This leads generally to poorer auditory discrimination and may contribute especially to reduced sensitivity to the temporal fine structure of sounds and to poor pitch perception. Hearing aids compensate to some extent for the effects of threshold elevation and loudness recruitment by the use of multi-channel amplitude compression, but they do not compensate for reduced frequency selectivity or loss of inner hair cells/synapses/neurons. The multi-channel compression processing used in hearing aids can impair some aspects of the perception of music, such as the ability to hear out one instrument or voice from a mixture. Systems for reducing acoustic feedback can also have undesirable side effects when listening to music. Finally, the limited frequency range and irregular frequency response of most hearing aids is associated with poor sound quality for music.

Josh H. McDermott
Massachusetts Institute of Technology

Computational auditory scene analysis as causal inference

A central computational challenge of everyday hearing is the need to separate the distinct causes of sound in the world. The most commonly discussed version of this problem occurs with concurrent sound sources, often termed the ‘cocktail party problem’. However, analogous problems are posed by reverberation, in which the sound from a source interacts with the environment (via reflections) on its way to the ears, as well as by sound-generating object interactions, in which the physical properties of multiple objects jointly determine the sound. Dating back to Helmholtz, perceptual judgments have been considered the result of unconscious inference, in which our perceptual systems determine the most likely causes of sensory stimuli in terms of structures and events in the world. Despite the conceptual appeal of this view, perceptual inference has historically been difficult to instantiate in working computational systems for all but the simplest perceptual judgments. In this talk I will revisit the notion of scene analysis as inference, leveraging recent computational developments that make inference newly feasible and exploring neglected classes of everyday scene analysis problems along with classical auditory scene analysis.

Mouny Elhilali
Johns Hopkins University

Attention at the cocktail party: cognitive control of auditory processing

In our daily lives, we are constantly challenged to attend to specific sound sources or follow particular conversations in the midst of competing background chatter - a phenomenon referred to as the ‘cocktail party problem’. Faced with this constant challenge, the auditory system exhibits a remarkable ability to adapt to its listening environment, driven by attentional control signals that bias how we process incoming information from our surrounds. These feedback mechanisms reshape the encoding of sensory information in the brain. Recent findings have been amending our views of processing in the auditory system; replacing the conventional view of ‘static’ processing in sensory cortex with a more ‘active’ and malleable mapping that rapidly adapts to the task at hand and listening conditions. After all, humans and most animals are not specialists, but generalists whose perception is shaped by experience, context and changing behavioral demands. The talk will discuss lessons learned from biology with regard to adapting to an ever-changing acoustic environment and the impact on building truly intelligent audio processing systems.

SPECIAL GUEST

Robert J. Zatorre
Montreal Neurological Institute, McGill University

Musicians at the cocktail party: neural correlates of top-down and bottom-up segregation mechanisms

Segregating sound mixtures makes demands on multiple cognitive and neural mechanisms that musical training may enhance or exploit. In a series of studies we have documented the music-related enhancement behaviorally in the context of speech in noise, and also in a selective attention context with competing speech streams. Using functional MRI, we observed that music-related enhanced speech-in-noise perception was associated with better decoding of speech in auditory areas at high signal-to-noise ratios (SNR), whereas under low SNR conditions the enhancement was associated with decoding in frontal and motor cortical regions. We interpret this finding as indicating a shift from bottom-up to top-down mechanisms depending on the quality of the input, with musicians being better able to deploy either mechanism as a function of the conditions. We then used MEG to look at the neural representation of competing speech streams via decoding of the neural signature (amplitude envelope) of attended vs unattended items. The behavioral advantage associated with musical training was related to enhanced ability to represent both streams in auditory cortex, consistent with their capacity to follow multiple sound streams in music. These cognitive neuroscience approaches help us to develop better models to explain why musicians are good at cocktail parties (apart from their reputed drinking abilities).

Thursday, March 28, 2019

Session 1: Hearing loss in music and speech (Chair: Kai Siedenburg)

Using the Adaptive Music Perception (AMP) test to understand the effects of hearing loss and hearing aids on the perception of musical scenes

Frank A. Russo, *Ryerson University*

It stands to reason that hearing loss may have some influence over the perception of musical scenes. However, there are few standardized tests available for measuring this objectively. The Adaptive Music Perception (AMP) test (Kirchberger & Russo, 2015) was designed to adaptively measure discrimination thresholds for the following basic dimensions of music: pitch, level, duration, dissonance, intonation, spectral brightness, attack time, and spectral irregularity, using a two-alternative forced choice paradigm and adaptive threshold testing. In our research using this test to date, we have assessed discrimination thresholds in participants with and without hearing loss, including those who are experienced hearing aid users. Results show that participants with mild-to-moderate hearing loss who do not use hearing aids have worse discrimination thresholds than those with normal hearing for pitch, harmony, spectral brightness and spectral irregularity. Moreover, participants with mild-to-moderate hearing loss who are experienced hearing aid users show no improvements in discrimination thresholds while using their hearing aids; in fact, the thresholds obtained for level and attack time are worse while using hearing aids. These findings suggest that perception of musical scenes is adversely affected by hearing loss, and that for those with mild-to-moderate hearing loss, current hearing aids to not help.

How many factors do we need to describe and model suprathreshold "distortion"?

Birger Kollmeier, *University of Oldenburg*

An overview of some of the work in our Cluster of Excellence Hearing4all (Oldenburg/ Hannover) is given with a focus on recognizing speech in noise – the classical “Cocktail Party Problem” which becomes more acute with increasing hearing loss and age. The Framework for Auditory Discrimination Experiments (FADE, Schädler et al., JASA 2016) is employed for predicting patient performance employing the German Matrix sentence test. Using a simple approach to include suprathreshold performance deficits in the model (Kollmeier et al., 2017), several of the “classical” auditory-model-based assumptions and theories about hearing loss can be tested: For example, the negative effect of increased auditory filterwidths in hearing-impaired listeners is much less pronounced than currently believed. Also, a very simple combination of a filterbank frontend with some amplitude modulation filter properties can already explain the observed speech recognition thresholds in strongly fluctuating noise much better than current standard models (SII, ESII, STOI, mr-sEPSM, Schubotz et al., 2016, Spille et al., 2018). Hence, the number of factors to describe suprathreshold “distortion” or loss in information processing – especially with relation to auditory scene analysis – seems to be smaller than “classically” assumed. There is still much to learn about our ears – and machines definitely help!

Session 2: Advances in hearing devices (Chair: Marc Réne Schädler)

Binaural speech enhancement and cue preservation algorithms

Simon Doclo, *University of Oldenburg*

Despite the progress in speech enhancement algorithms, speech understanding in adverse acoustic environments with background noise, competing speakers and reverberation is still a major challenge for many hearing aid users. In this presentation, some recent advances in multi-microphone noise reduction algorithms for binaural hearing aids will be presented. The objective of these algorithms is not only to selectively extract the target speaker and to suppress background noise and reverberation, but also to preserve the auditory impression of the acoustic scene. Aiming at preserving the binaural cues of all sound sources while not degrading the noise reduction performance, different extensions of the binaural minimum variance distortionless response (MVDR) beamformer and the binaural multi-channel Wiener filter will be presented, both for diffuse noise as well as for interfering sources. These extensions aim at either preserving the interaural time and level difference cues of the interfering sources or the interaural coherence of the background noise, based on psycho-acoustically motivated boundaries. Evaluation results will be presented in terms of objective performance measures as well as subjective listening scores for speech intelligibility and spatial quality. Furthermore, first steps towards incorporating computational acoustic scene analysis and external microphones into these algorithms will be discussed.

Compensation for impaired auditory scene analysis

Niels H. Pontoppidan, *Eriksholm Research Centre, Oticon A/S*

People with hearing problems find situations with competing voices rather difficult. In fact, the presence of two competing voices limit their ability to segregate the two and to tune in and attend to just one. Previous research reported that suprathreshold hearing deficits are indeed an important part of the problem and found correlations between identifying simple auditory contrasts and listening tests where one speech signal was the target disturbed by another speech masker. Similar correlations have been reported for competing voices situations where the listener is not initially aware of which of the speech signals that is the target and masker. These problems affect people with sensorineural hearing loss as well as deaf people that need cochlear implants to understand speech at all. In addition, these hearing problems also seem to play into the ability to perceive and enjoy music. Recently, by applying speech separation with deep neural networks to mixtures of known competing voices, we have demonstrated that this separation increases speech recognition and enhances segregation for listeners with hearing impairment. Before such technology becomes available to people with hearing impairment, the robustness with respect to gradual and sudden voice changes, reverberation, and other noises must be investigated and improved.

Multi-modal hearing devices and their evaluation

Volker Hohmann & Giso Grimm, *University of Oldenburg*

Spatial filtering and decomposition of sounds into acoustic source objects is increasingly investigated for speech enhancement in hearing aids. However, with increasing performance and availability of these 'space aware' hearing aid algorithms, knowledge of the user's personal listening preferences and knowledge of the attended source becomes crucial. Here we present an algorithm, which combines information of the subject's eye gaze with an acoustic analysis of the sound source positions to identify the attended source from a mixture of sources. Gaze direction is recorded by electrooculography (EOG) combined with a head tracking system, which could in principle be integrated in hearing aids. The spatio-temporal distribution of source positions is estimated from input signals to a binaural hearing device. Closed-loop hearing devices that use (electro-)physiologic sensor signals to estimate attention and behaviour require "subject-in-the-loop" evaluation methods, which simulate realistic interactive environments to reproduce natural behaviour. This means that audio-visual signals are required that also include behavioural simulation of the sources, e.g., simulated lip movement, or simulated conversational gaze behaviour. [Project funded by DFG SFB 1330 project B1.]

Session 3: Stream formation (Chair: Emily Coffey)

A comparative view on auditory scene analysis and informational masking

Georg M. Klump, *Oldenburg University*

Segregating sound sources in a complex acoustic environment facilitates the analysis of signals from a specific source. The benefits of assigning sounds to specific sources accrue to all species communicating acoustically. Thus, the ability for auditory scene analysis is widespread among different animals. Animal studies allow for a deeper insight into the neuronal mechanisms underlying auditory scene analysis. A number of stimulus paradigms are suitable for probing the mechanisms reflecting different levels of integration, i.e., by evaluating perception of human subjects and animals and relating the data on perceptual sensitivity to the sensitivity evident in responses of neurons in the auditory pathway. A range of stimulation paradigms that have been extensively applied in both psychophysical and neurophysiological studies relies on auditory stream segregation based on salient cues that allow distinguishing the sounds from different sources. Auditory stream segregation not only affects the perception of the temporal pattern of sound sequences, it also has an impact on the discrimination of signals in the context of informational masking. Examples are provided with such stimulus paradigms being applied in human and animal psychophysical studies and in studies of populations of auditory neurons.

Auditory streaming cues in cochlear implant listeners

Jeremy Marozeau, *Technical University of Denmark*

The cochlear implant (CI) is the world's first successful surgically implanted sensory prosthesis. It produces a sound sensation by stimulating the hearing nerve directly with electrical impulses from electrodes implanted in the cochlea. However, it does not restore normal hearing. CI listeners have great difficulties with understanding speech in a noisy environment or to appreciate complex polyphonic music. Both of those problems are related to the ability to segregate auditory stream. Many studies have investigated the principles or cues that allow the healthy auditory system to perceptually group sounds into streams and selectively attend to one of them. However, the number of studies investigating these processes in CI listeners is limited, and their findings are contradictory. In this presentation, I will review some of those studies and present some recent data that showed that perceptual differences elicited by varying either the place or the pulse rate of the electrical stimulation can allow CI listeners to group the sounds into auditory streams.

Session 4: Multistability (Chair: Steven van de Par)

Multistability in auditory scene analysis

Alexandra Bendixen, *Chemnitz University of Technology*

Auditory scene analysis (ASA) is a prime example of perceptual multistability: When exposed to an ambiguous sound sequence, listeners' perception switches back and forth between different interpretations. Such multistability allows studying ASA processes without confounding changes in stimulus input. Along these lines, much insight into ASA has been gained by using the ‚ABA_' auditory streaming paradigm as a simplified model scenario with two potential sound sources. We extend this approach towards more complex source configurations. In subjective-report procedures, we show that listeners spontaneously report hearing more than one sound source in the perceptual foreground at once. To examine the validity of listeners' subjective reports, we develop audio-visual stimulus protocols that allow us to collect correlates of auditory foreground formation via eye-tracking. The eye-tracking results are in good agreement with the indicated auditory percepts, suggesting that foreground perception is validly reported. To study how the perceptual background is structured in complex auditory scenes, we combine behavioral measures with electroencephalography (EEG). Our results suggest that different representations of the perceptual background are held in parallel. Implications of these findings for current ASA models will be discussed.

Ambiguity of spectral shift judgments: a case for spectral envelopes

Kai Siedenburg, University of Oldenburg

Recent research has described strong effects of prior context on the perception of ambiguous half-octave pitch shifts of Shepard tones (Chambers et al., Nat. Commun 2017). In this talk, I will describe similar effects for spectral envelope shifts. Specifically, I will describe a set of experiments based on brightness shift judgments of harmonic tone complexes with cyclic spectral envelope components and fixed fundamental frequency. Experimental results indicate that frequency shifts of the envelopes are perceived as systematic shifts of brightness. Analogous to the work of Chambers et al., the perceptual ambiguity of half-octave shifts resolves with the presentation of prior context tones, which constitutes a novel context effect for the perceptual processing of spectral envelope shifts. Ongoing research investigates the dominance relations between concurrent shifts of spectral fine structure and envelope with a particular focus on the role of the (in)harmonicity of the spectral fine structure. By generalizing the classic Shepard tone stimulus, this work attempts to shed new light on the interplay of three key elements of auditory scene analysis: pitch, timbre, and harmonicity.

Binding in auditory scene analysis: Insights from behavioral and pupillometric experiments

Jackson Graves & Daniel Pressnitzer, *École Normale Supérieure, PSL University Paris*

One basic process of auditory scene analysis is to bind together distinct frequency components, over time and frequency, to represent individual sound sources. One interesting case is when the binding over time is ambiguous, such as when two Shepard tones are presented with a half-octave interval. The perceived direction of the resulting pitch-shift is variable and strongly affected by context, even though listeners are mostly unaware of the stimulus ambiguity. In two new experiments, we investigated potential implementation levels of the feature binding process that we hypothesize is at work in this phenomenon. First, a monaural test interval of two Shepard tones was preceded by opposing context sequences in the two ears. We observed effects of both attended ear and ear of entry, suggesting that feature binding is amenable to both attention and spatial location. Second, listeners' pupil size was measured while they judged the change direction of Shepard tone intervals. We observed increased pupil dilation for conditions of ambiguous change direction, suggesting that ambiguity was represented within the auditory system even when listeners showed no behavioral awareness of it. Together, the results are compatible with the idea that auditory feature binding occurs at multiple levels of auditory processing.

Friday, March 29, 2019

Session 5: Attention (Chair: Kai Siedenburg)

Attentional capture in auditory scenes: salient events shape perception by improving the neural representation of the capturing stream

Jannis Hildebrandt, *University of Oldenburg*

A sudden, salient auditory stimulus may capture attention towards a previously unattended auditory object in multi-source auditory scenes. Such a shift of attention may not only serve to allocate cognitive resources to a potentially relevant signal, but also change sensory processing of the capturing object. Surprisingly little is known about effect of exogenously triggered attention on auditory processing. We aimed to characterize both perceptual consequences and underlying neural mechanisms following attentional capture. We devised a stimulus evoking a clear two-stream percept, mimicking distinct sources, and tested how salient, attention-capturing cues in either of the two streams affected detection of targets. We tested our paradigm in human psychophysics experiments, confirmed the behavioral results in mice in a corresponding go-no go paradigm, and performed electrophysiological recordings in the primary auditory cortex (ACx) of awake-behaving mice. In both human and mouse psychophysics capturing attention improved auditory temporal resolution. Recordings from populations of single-units in ACx revealed that valid cues improved target representation while invalid cues obscured target encoding. Notably, modulation was strongest for a suppressive component of the responses that may signal prediction. Thus, valid and invalid cues could affect perception by modulating predictive signals to either of the two streams.

Auditory selective attention in noisy environments – differences between children and young adults

Sophie Nolden*, Karin Loh, Edina Fintor, & Janina Fels, **Goethe University Frankfurt*

We developed an experimental paradigm to test cognitive flexibility of auditory attention and susceptibility to noise in children (6-10 years old). The task was inspired by previous work with young and older adults. Participants focused on one of two binaurally presented voices in order to perform a two-choice classification task. Target and distractor voice appeared on two of four possible locations. The target voice could change from trial to trial. In addition, participants were sometimes exposed to background noise with the frequency spectrum of children's voices. Twenty-four children and 24 young adults participated in the auditory attention and listening task. Their anthropometric sizes were considered for an aurally accurate reproduction of the acoustic scene. Children and young adults showed similar patterns of auditory attentional control. However, children showed a different pattern of results compared to young adults when confronted with noise. The current study increased our knowledge of auditory attention in children in noisy educational buildings.

Session 6: Musical scenes I (Chair: Stefan Weinzierl)

Scene analysis as a psychological foundation for musical orchestration

Stephen McAdams, *McGill University*

Music uses the qualities of different instruments to create specific perceptual, expressive, and emotional effects that composers sculpt over time. Timbre perception is multifaceted and contributes in many ways to the perceptual organization of musical structures. Orchestration taken its broadest sense is the conception, selection, combination, and juxtaposition of sound qualities to achieve a specific musical aim. The aim of this work is to develop a theoretical ground for orchestration practice starting with the structuring role that timbre can play in music, using scene analysis as a starting point. Many facets of musical structuring are achieved by auditory scene analysis, the perceptual grouping processes that: 1) fuse different acoustic components into events (e.g., instrumental blend), 2) integrate events into one or more auditory streams or other sequential groupings (e.g., surface textures or orchestral layers), 3) segment groups of events into motifs, phrases, and sections (e.g., antiphonal contrasts, section boundaries), and 4) form larger-scale units encompassing changes in orchestration that are extended over time (e.g., orchestral gestures). The roles that timbre plays in the manifestation of these principles in orchestration practice will be considered as a point of departure for developing a perceptually based theory of orchestration.

A review of sound engineers' advice for manipulating complex soundscapes

Trevor R. Agus, *Queen's University Belfast*

When researching auditory scene analysis, there is often a trade-off between control of the sound scene and a level of acoustical complexity that is representative of our everyday soundscapes. Sound engineers work almost exclusively with paradigms that combine multiple, complex stimuli with reasonably controlled conditions. They develop practical knowledge that allows them to manage the audibility, clarity and blend of the soundscapes they curate. Here, I review the

documented observations and strategies of sound engineers in terms of their implications for the psychoacoustics of complex soundscapes. As the tools of sound engineers generally provide control over levels and spectra, it is unsurprising that many observations implicate loudness, masking and timbre. Clarity of sounds in general, not just speech, is discussed regularly, and blend and distance are also recurring themes. The preservation of transients is also considered important by several of the sound engineers. Dynamic-range management is claimed to have implications for many of the percepts discussed, including those relating to auditory scene analysis. The sound engineers' experience with dynamic-range compression seems pertinent to the design of audio prostheses, which also often use compression. Implications for mixing and mastering audio for hearing-impaired listeners are discussed.

Analyzing assimilated musical scenes: sound mass perception and its acoustical correlates

Jason Noble, Max Henry, Étienne Thoret, & Stephen McAdams, *McGill University*

Traditional musical categories such as harmony, melody, and rhythm are predicated on intervallic relations between discrete units, but contemporary music sometimes deliberately subverts the listener's ability to segment the auditory scene, turning attention instead to the properties of the massed musical or sonic totality. Albert Bregman (1990) likened such sound masses to "the spraying shower or the flowing hair in vision," defined in perception by their global "granularity" rather than individual lines. Sound mass composition exploits the same perceptual principles that underpin the streaming of independent musical lines in polyphonic music (as detailed by David Huron, Mus. Perc. 2001), with the opposite perceptual effect of assimilation resulting from radically different musical organization. While many authors have described the technical methods and artistic aims of sound mass composers, little empirical research has studied the extent of agreement between listeners as to what constitutes a mass, or the acoustical properties that lead to sound mass integration. This presentation reports results from several experiments confirming significant coherence between listeners in their perceptions of sound mass. Acoustical analyses including classical descriptor multilinear regression, modeling of cochlear and mid-level processing, and neuromimetic representations illuminate properties of the acoustical signal that correlate with participants' ratings from these experiments.

Session 7: Musical scenes II (Chair: Henning Schepker)

Making music more accessible for cochlear implant listeners

Waldo Nogueira, *Hannover Medical School*

Cochlear implants (CIs) have become remarkably successful in restoring the hearing abilities of profoundly hearing-impaired people. Although in most cases speech understanding with CIs reaches around 90%, key musical features such as pitch and timbre are poorly transmitted by CIs, leading to a severely distorted perception of music. Because music is a ubiquitous means of sociocultural interaction, this handicap significantly degrades the quality of life of CI users. Therefore, in this contribution, recent developments that enable CI users to access music are presented. More concrete the limitations pitch and timbre perception with CIs are presented as well as its implications for music perception. Next different emerging strategies for improving CI users' music enjoyment, such as customized music compositions, music pre-processing methods for the reduction of signal complexity, and improved sound coding strategies will be reviewed together with subjective and objective instrumental evaluation procedures.

Time dilations in music and audiovisual media

Clemens Wöllner, *University of Hamburg*

Music as an ephemeral form of art that works simultaneously on different hierarchical levels may alter our sense of time and space. This talk will focus on two recent studies that analyzed the impact of music on perceived time. First, using a sensorimotor synchronization paradigm, we investigated whether tapping with different metrical levels in music (half notes, quarter notes, eighth notes) influences perceived duration. Results provide evidence for temporal dilations in relation to attended event density. Second, we investigated slow-motion videos that are typically underscored by emotionally expressive music. Based on theories that suggest links between distorted time experiences and increased cognitive as well as physiological processes, effects were found on perceived emotion, peripheral autonomic arousal (GSR, HR, RSP, pupillary dilations) and eye movements. It is argued that time perception, attentional processes and individual experiences are essential for the integration and segregation of perceived scenes.

Session 8: Are musicians special? (Chair: Kirsten Wagener)

Incidence and relative risk of hearing disorders in professional musicians

Gunter Kreutz*, Veronika Busch, Iris Pigeot, Wolfgang Ahrens & Tania Schink, **University of Oldenburg*

Hearing disorders have been associated with occupational exposure to music. Musicians may benefit from non-amplified and low-intensity music but may also have high risks of music-induced hearing loss. The goal of this study was to compare the incidence of hearing loss (HL) and its sub-entities in professional musicians with that in the general population. We performed a historical cohort study among insurants between 19 and 66 years who were employed subject to social insurance contributions. The study was conducted with data from three German statutory health insurance providers covering the years 2004–2008 with about 7 million insurants. Incidence rates with 95% CIs of HL and the sub-entities noise induced hearing loss (NIHL), conductive HL, sensorineural HL, conductive and sensorineural HL, as well as tinnitus were estimated stratified by age, sex and federal state. A Cox regression analysis was conducted to estimate adjusted HRs and two-sided 95% CIs for HL and its sub-entities. More than 3 million insurants were eligible, of whom 2,227 were identified as professional musicians (0.07%). During the 4-year observation period, 283,697 cases of HL were seen, 238 of them among professional musicians (0.08%), leading to an unadjusted incidence rate ratio of 1.27. The adjusted hazard ratio of musicians was 1.45 (95% CI 1.28 to 1.65) for HL and 3.61 (95% CI 1.81 to 7.20) for NIHL. Hence, professional musicians have a high risk of contracting hearing disorders. The use of already available prevention measures should reduce the incidence of HL in professional musicians.

Speech perception is similar for musicians and non-musicians under wide range of conditions

Sara M. K. Madsen, *Technical University of Denmark*

It remains unclear whether musical training is associated with improved speech understanding in a noisy environment, with different studies reaching differing conclusions. Even in those studies that have reported an advantage for highly trained musicians, it is not known whether the benefits measured in laboratory tests extend to more ecologically valid situations. This study aimed to establish whether musicians are better than non-musicians at understanding speech in a background of competing speakers or speech-shaped noise under more realistic conditions, involving sounds presented in space via a spherical array of 64 loudspeakers, rather than over headphones, with and without simulated room reverberation. The study also included experiments testing fundamental frequency discrimination limens (F0DLs), interaural time differences limens (ITDLs), and attentive tracking. Sixty-four participants (32 non-musicians and 32 musicians) were tested, with the two groups matched in age, sex, and IQ as assessed with Ravens Advanced Progressive matrices. There was a significant benefit of musicianship for F0DLs, ITDLs, and attentive tracking. However, speech scores were not significantly different between the two groups. Taken together with the results from other studies, the results suggest no musician advantage for understanding speech in noise under a variety of conditions.

Hearing-in-noise perception: why are musicians better?

Emily Coffey, *Concordia University*

To perceive speech in a naturalistic environment a listener might rely on the accuracy with which their brain encodes sound, but also use their familiarity with the speakers voice, visual and spatial cues, and predictive cues. Musicians seem to be better at HIN perception, but results from previous work are mixed. Are musicians really better? And if so, is it related to use of such cues? In a recent review, we systematized studies addressing musicianship enhancement on HIN perception. We found that the majority of studies did not include visual and spatial cues, and in the auditory paradigms, a pattern emerged: a musician advantage was found more often when a restricted set of cues was offered. We hypothesized that musicians' superior encoding of fine acoustic cues partly underlies HIN perception, a result that was supported by an MEG study in which periodicity encoding throughout the auditory system was related to HIN performace. However, other studies have also found musician advantages in audio-visual integration, spatial auditory perception, selective attention, and working memory, which are all relevant to HIN perception. To address these issues, we designed and validated a new task: the Music-In-Noise Task (MINT), which is a tool for dissecting complex auditory perception. Musicians scored particularly well on subtasks involving visual, spatial and predictable cues. The task materials are offered to the community for further experimentation and development, and we have conducted several studies using EEG and fMRI that further address the value of these cues in HIN perception.

This symposium is made possible by the Hanse Wissenschaftskolleg Associated Junior Fellowship Program and the German Research Foundation (DFG) International Conference Program.