



21st Annual Auditory Perception, Cognition, & Action Meeting

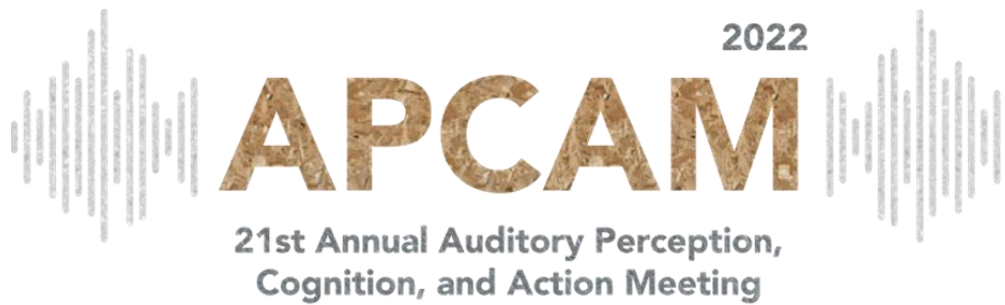


Thursday, November 17th

Sheraton Boston Hotel

Boston, MA

8:00am - 5:00pm



Welcome Letter

Welcome to the 21st annual Auditory Perception, Cognition, and Action Meeting (APCAM 2022)! We are excited to host an in-person meeting for the first time since 2019. Since its founding in 2002, APCAM's mission has been "...to bring together researchers from various theoretical perspectives to present focused research on auditory cognition, perception, and aurally guided action." We believe APCAM is a unique meeting that blends both basic and applied research from different theoretical perspectives and processing accounts using numerous types of auditory stimuli (including speech, music, and environmental noises). As noted in previous programs, the fact that APCAM continues to flourish is a testament to the openness of its attendees to consider multiple perspectives and value diversity, which is a principle characteristic of scientific progress.

APCAM is affiliated with the journal *Auditory Perception and Cognition (AP&C)*, which features both traditional and open-access publication options. Presentations at APCAM 2022 are eligible to submit a brief report for consideration in a special issue of *AP&C*. If you did not indicate a preference to be considered for the special issue when you submitted your abstract to APCAM, you can contact the editors of *AP&C*, Michael Hall (hallmd@jmu.edu) and Mike Russell (mirussell@bellevue.edu), to ask for such consideration. In addition, we encourage you to submit your other work on auditory science to *AP&C*.

APCAM is also affiliated with the Auditory Perception and Cognition Society (APCS) (<https://apcsociety.org>). This non-profit foundation is charged with furthering research on all aspects of audition. The \$30 registration fee for APCAM provides a one-year membership for APCS, which includes an individual subscription to *AP&C* and reduced open-access fees for publishing with *AP&C*.

As an affiliate meeting of the 63rd Annual Meeting of the Psychonomic Society, APCAM is indebted to the Psychonomic Society for material support. We acknowledge and are grateful for their support, and we ask that you pass along your appreciation to the Psychonomic Society for their continued support of APCAM.

We appreciate all our colleagues who contributed to this year's program. We thank you for choosing to share your work with us, and we hope you will continue to contribute to APCAM in the future. This year's meeting features a keynote presentation, *New Models of Human Hearing via Machine Learning*, by **Josh McDermott**; 17 spoken sessions; and 28 posters that cover a wide range of topics in auditory science. We are confident that everyone attending APCAM will find something interesting, relevant, and thought-provoking. If you use Twitter, you can tweet about APCAM using the hashtag: **#APCAM2022**

If there are issues that arise during the meeting, or if you have thoughts for enhancing future meetings, do not hesitate to contact any committee member. We wish you a pleasant and productive day at APCAM!

Sincerely,

The APCAM 2022 Organizing Committee

Timothy L. Hubbard (Chair)

J. Devin McAuley

Kathleen C. McCulloch

K. Jakob Patten

Peter Q. Pfordresher

Hannah Shatzer

Map

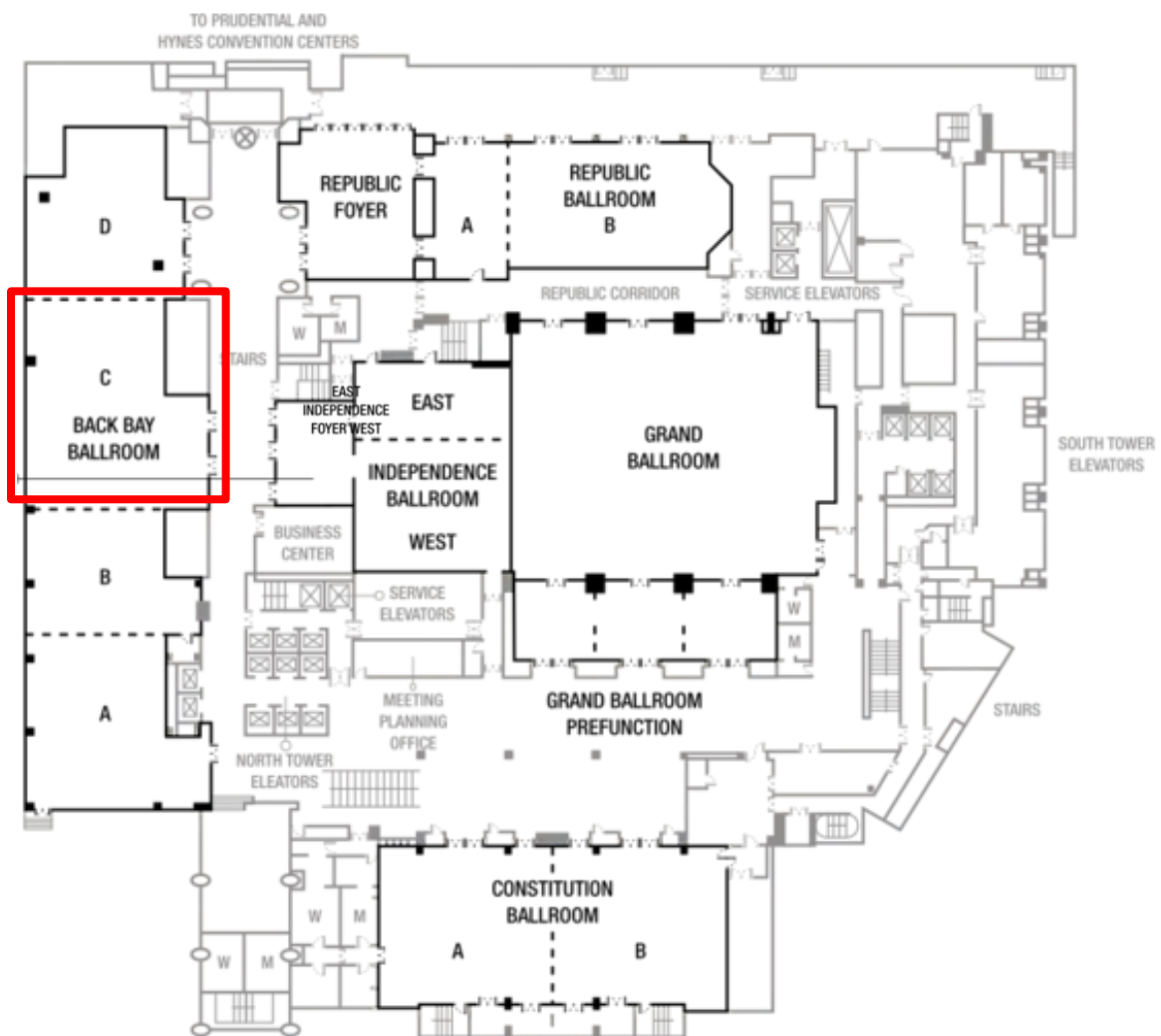
APCAM 2022 will be held in:

Back Bay C (2nd Floor)

Sheraton Boston Hotel

39 Dalton St.

Boston, MA 02199



Posters will be displayed in:

Hynes Convention Center Ballroom (3rd Floor)

900 Boylston St.

Boston, MA 02115



Full Schedule

8:00	Registration	
8:10	Welcoming Remarks	
8:15	Temporal Processing Talks	
8:15	Effect of stimulus type on temporal summation in young and aging adults	Leah Fostick, Ariel University
8:30	Spatial Contributions to Auditory-Motor Synchronization	Carolyn Kroger, Michigan State University
8:45	Analysis of auditory-related alpha oscillations (tau rhythms) in the EEG	Matthew G. Wisniewski, Kansas State University
9:00	Temporal representation influenced by auditory imagery and sensory information in working memory	Sharath Chandra Ramakrishnan, University of Texas at Dallas
9:15	Using mobile technologies for telerehabilitation in children with ADHD: Validation study of a rhythm-based serious game	Hugo Laflamme, University of Montreal
9:30	Membership Meeting	
10:00	Break	
10:15	Speech and Music Talks	
10:15	Neural correlates of multimodal speech perception in cochlear-implanted and normal-hearing adults	Hannah Shatzer, Toronto Metropolitan University
10:30	Learning mechanisms in phonetic cue weighting: What we learn from a single cue	Vsevolod Kapatsinski, University of Oregon
10:45	Distributional learning of a novel speech contrast by bilingual and monolingual listeners	Siyu Lin, Emory University
11:00	The role of rhythm complexity in cardiac dynamics during auditory perception and production	Shannon E. Wright, McGill University
11:15	Undetectable Very-Low Frequency Sound Increases	Michael Hove, Fitchburg

	Dancing at a Live Concert	State University
11:30	Poster Session (Hynes Ballroom, 3rd Floor)	
11:30	Poster set-up	
11:45	Odd-numbered authors at posters	
12:15	Even-numbered authors at posters	
1:00	Lunch	
2:15	Keynote: Josh McDermott, MIT <i>New Models of Human Hearing via Machine Learning</i>	
2:45	Auditory Memory and Attention Talks	
2:45	Acoustic change detection during associative auditory learning and explicit memory is related to enhanced signal detection at retrieval	Manda Fischer , University of Toronto; Rotman Research Institute
3:00	Perceived Similarity and Recognition Memory for a Set of Novel Auditory Stimuli	Nathan Gillespie , University at Albany, SUNY
3:15	Is Auditory Preemption a Thing? Five Reasons to be Skeptical	Michael A. Nees , Lafayette College
3:30	Triggering misophonia: The importance of spectral information, temporal information, and action identification	Sawas Kazazis , McGill University
3:45	Break	
4:00	Spatial Localization Talks	
4:00	The Auditory Looming Bias is Influenced by Visual Information	Maggie K. McCracken , University of Utah
4:15	A cue or a distractor? Automatic attention in spatial discrimination	Norbert Kopco , P.J. Safarik University
4:30	Inferring object interactions from sound	Vinayak Agarwal , MIT

Poster Session

1-4	Auditory Memory and Attention	
1	ERP correlates of auditory peak shifts in stimulus generalization	Chelsea Joyner , Kansas State University
2	Accuracy and Precision of Reproduced Short Term Memories	Michael A. Tollefsrud , Kansas State University
3	Exploring the Disruptive Role of Background Speech on Short-term Memory: What Role for Whispering?	Kathleen C. McCulloch , University of Central Lancashire
4	Cognitive control of auditory attentional capture: Examining the interaction of contextual factors and trait capacity for steadfast task-engagement	John E. Marsh , University of Central Lancashire
5-12	Auditory Scene Analysis	
5	Dimensionality of Natural Auditory Scene Perception: A Factor Analysis Study	Margaret A. McMullin , University of Nevada, Las Vegas
6	Auditory Judgments of Egocentric and Allocentric Motion and Motion Direction	Mike Russell , Bellevue University
7	Auditory Perception of Occlusion in the Absence of Visual Information	Mike Russell , Bellevue University
8	Streaming Sound Texture in Auditory Scenes	Jarrold M. Hicks , MIT
9	Predictive coding : Electrophysiological signature of timing predictions	Amour Simal , University of Montreal
10	Pleasantness altered by misidentification of everyday sounds	Laurie M. Heller , Carnegie Mellon
11	A preliminary evaluation of potential perceptual contributions from the skull	Laura Reinert , James Madison University
12	Categorization of environmental sounds according to the manner of contact between solids	Benjamin H. Stone , James Madison University

13-18	Music Perception	
13	Time, intensity, and frequency contributions to clasp perception in cyclic musical rhythms	Benjamin H. Stone, James Madison University
14	Auditory Statistical Learning in Classical and Jazz Musicians	Erica Knowles, Berklee College of Music
15	The Stability of the Speech-to-Song Illusion: Individual Differences	Rodica Constantine, University of Nevada, Las Vegas
16	Misophonic Experience and Musicality	Rodica Constantine, University of Nevada, Las Vegas
17	Are Melodies a Useful Mnemonic Cue for Word List Recall?	Laura M. Getz, University of San Diego
18	Perceptual evaluation of a novel method for synthesizing musical sounds	Michael D. Hall, James Madison University
19-22	Cross-Modal Perception	
19	Distance perception of objects using visual-to-auditory sensory substitution: comparison of conversion methods based on sound intensity and envelope modulation	Camille Bordeau, Université Bourgogne Franche-Comté
20	Cochlear implant users and normal-hearing listeners have contrasting hemodynamic changes during auditory beat processing despite similar behavioral response	Samantha R. O'Connell, University of Southern California
21	Restrictive eating tendencies and variability in audiovisual integration	Annika Beatteay, University of New Brunswick
22	Behavioral Markers of Age-Related Hearing Loss Revealed through the Attentional Network Task	Gennadiy Gurariy, Medical College of Wisconsin; Marquette University

23-28	Speech Perception and Production	
23	Acoustic profile of French emotional sentences uttered with French and Quebecois accents	Amélie Rolinat , University of Montreal
24	Voice-induced synesthesia, emotion and plasticity	Cathy Lebeau , UQÀM
25	It's not who you are, or what you say, but how many people say it: A multi-talker study of vocal confidence	Katie Asaro , Kent State University
26	Sensorimotor Processing of Vocal Emotion	Emma Greenspon , Monmouth University
27	An acoustic analysis of probabilistic language used during COVID-19 White House Press Briefings	Yin Zhang , Kent State University
28	Assessing the Effects of "Native Speaker" Status on Classic Findings in Speech Perception	Julia Strand , Carleton College

Talk Abstracts

Temporal Processing (8:15-9:15)

8:15 **Effect of stimulus type on temporal summation in young and aging adults**

Leah Fostick*
Chanit Cohen

Ariel University
Ariel University

Aging adults have higher hearing thresholds. They also have difficulty perceiving short sounds. Temporal summation, a neurophysiological principle, combines loudness and duration and states that longer sounds are perceived as louder than shorter ones. Therefore, increasing the duration of sounds may help aging adults to hear better. However, it is not clear whether temporal summation, a temporal-processing-based mechanism, will be as effective among aging adults that show difficulties in temporal processing. Also, most studies tested temporal summation for pure tones (PT) and showed it to be effective mainly in lower frequencies. Therefore, it is unclear whether temporal summation will be effective in complex sounds we hear daily.

Temporal summation was tested on 24 normal-hearing young adults (age 20-35 years) and 24 aging adults (age 60-75 years) for five speech-shape-noise (SSN) sounds (of the speech sounds /a/, /i/, /u/, /sh/, and /m/), three PTs (500, 1000 and 4000 Hz), and for white noise (WN). Hearing thresholds in each sound were tested for five different durations (1, 5, 20, 50, and 100 ms). A general decline in hearing thresholds was observed as stimulus duration increased, reflecting a temporal summation for all stimuli, but for /m/ (for all participants) and /sh/ (for aging adults only). In line with the literature, the largest slope was 500 and 1000Hz pure tones than temporal summation for complex sounds. Aging adults' temporal summation was larger than young adults' for WN and /i/, but smaller for the high-frequency stimuli 1kHz, 4kHz, and /sh/.

These findings suggest that temporal summation is mostly effective for low-frequency PT but not less so for complex sounds. Also, the benefit aging adults can have from temporal summation in daily life is limited.

Author e-mail: leah.fostick@ariel.ac.il

[Back to Schedule](#)

8:30 Spatial Contributions to Auditory-Motor Synchronization

Carolyn Kroger*

Michigan State University

J. Devin McAuley

Michigan State University

Coordinating rhythmic movements with an auditory stimulus allows us to accomplish many different types of tasks, from playing music with a metronome to marching in synchrony with a drum. One method widely used to investigate processes underpinning the timing of rhythmic behavior is the synchronization-continuation tapping task. In this task, individuals synchronize finger taps with an auditory pacing signal then continue tapping at the same tempo once the metronome stops. Researchers using this method have made inferences about timing processes based on produced sequences of tap onsets, ignoring the continuous nature of movements and potential contributions of spatial elements of movements to synchronization-continuation tapping performance. In two motion-tracking experiments, we considered continuous movement dynamics during synchronization-continuation tapping to test predictions of an amplitude control hypothesis. This hypothesis proposes that people control the amplitude of their finger tapping movements in order to produce a target inter-tap interval (ITI). In the first experiment, participants performed an auditory synchronization-continuation tapping task for target ITIs ranging from 250 to 1150 ms. Results showed support for the amplitude control hypothesis: people changed tap amplitude as a function of produced ITI where amplitudes were lower for faster compared to slower tapping. In the second experiment, participants tapped at target amplitudes ranging from low to high and at target ITIs ranging from 250 to 1150 ms. Results showed that timing variability was greater for fast target ITI, especially for high target amplitudes. There was a tradeoff between ITI accuracy and amplitude accuracy where amplitude was lower for fast target ITI, suggesting that people reduced amplitude to produce fast target ITI at the cost of matching target amplitude. These experiments highlight the importance of considering spatial dynamics of movement to understand processes involved in timing rhythmic movements. Implications for understanding timing deficits in clinical populations will be discussed.

Author e-mail: krogerca@msu.edu

[Back to Schedule](#)

8:45 Analysis of auditory-related alpha oscillations (tau rhythms) in the EEG

Matthew G. Wisniewski*	Kent State University
Chelsea Joyner	Kent State University
Alexandria C. Zakrzewski	Kent State University
Scott Makeig	University of California, San Diego

Tau rhythms are characterized by sound responsive alpha band (~8 - 13 Hz) oscillations generated largely within auditory areas of the superior temporal gyri. Most research on tau has employed MEG or intracranial methods because of tau's elusiveness in the EEG. We conducted two experiments that develop a measure of tau from independent components analysis (ICA) decomposition of high-density EEG data. In Experiment 1, listeners were passively presented complex sounds while the EEG was recorded. Listeners' data was split into parallel processing pipelines entailing various filtering and ICA model choices with the goal of optimizing processing parameters. Analyses revealed that: 1) proper filtering and ICA model choice can yield a high percentage of sample subjects (~94%) showing a tau-related independent component (IC) localized in or near the superior temporal gyri, 2) the sound responsive dynamics of tau ICs parallel those observed with MEG and intracranial methods, and 3) tau dynamics are independent of stronger alpha rhythms in the brain (e.g., somatomotor mu, occipital alpha). Experiment 2 examined whether tau IC dynamics were impacted by attentional demands. Subjects performed an audio-visual spatial attention task in which they were instructed to respond according to auditory or visual stimuli. Tau ICs showed stronger alpha band (8-13 Hz) suppression during attend-auditory than attend-visual trials. Occipital alpha ICs showed the opposite. We expect that adoption of the ICA approach to EEG analysis will increase the rate and range of discoveries related to tau rhythms. Examinations of tau ICs may be especially useful in characterizing fluctuations in auditory attentional states.

Author e-mail: mgwisniewski@ksu.edu

[Back to Schedule](#)

9:00 Temporal representation influenced by auditory imagery and sensory information in working memory

Sharath Chandra Ramakrishnan* University of Texas at Dallas

Cochlear implant (CI) users cannot use auditory displays like notifications or ringtones since CI processing lacks fine harmonic resolution necessary for pitch perception. However, since CI performance is same as normal listeners in tempo and rhythm discrimination tasks, we designed an auditory display for use in CIs using a temporal auditory image, consisting of a series of variably pulsating tones using the phenomenon of beating between two interfering tones. Synthesis used a noise-excited envelope vocoder to simulate CI processing. Normal hearing volunteers performed a magnitude estimation task after training to retain the temporal variation of the auditory image across a sliding visual scale with reasonable accuracy (regression slope = 0.80, SE = 0.039). During training and recall, participants reported vocalizing the rate of beating or alluding to the noise vocoded sounds using auditory imagery from past experience. For instance, participants imagined increasing rates of pulsing from 32 to 100 Bpm as wheezing, slow jogging and panting, 200 - 350 Bpm as the sound of a salt shaker, or a hacksaw blade cutting wood, and higher rates until 600 Bpm to a helicopter or locomotive engine. These observations confirm the ability of motored subvocal speech sound production and covert verbal labeling of environmental stimuli, as a means to link sequences that can assist in the storage of rapidly alternating acoustic signals (Schulze et al., 2012). During magnitude estimation, participants slightly underestimated the slider position at higher rates of audible pulsing and found it harder to discriminate at rates of about 450 BPM and upwards. This points to possible imprecision in the sensori-motor synchronization provided by working memory, that enables one to update temporal intervals governing musical pulse sequences (Baddeley and Hitch, 1974). The results support growing evidence that auditory imagery is fortified by the maintenance of sensory information in working memory.

Author e-mail: SharathChandraRam@utdallas.edu

[Back to Schedule](#)

**9:15 Using mobile technologies for telerehabilitation in children with ADHD:
Validation study of a rhythm-based serious game**

Hugo Laflamme*	Int'l Lab. for Brain, Music and Sounds Research (BRAMS); Univ. of Montreal; Centre for Research on Brain, Language and Music (CRBLM)
Kevin Jamey	BRAMS; University of Montreal; CRBLM
Nick E. V. Foster	BRAMS; University of Montreal; CRBLM
Simon Rigoulot	BRAMS; University of Quebec at Trois-Rivières; CRBLM
Sarah Lippé	BRAMS; University of Montreal
Sonja A. Kotz	BRAMS; Maastricht University
Simone Dalla Bella	BRAMS; University of Montreal; CRBLM

Synchronization of movement to the beat of music involves monitoring one's own actions while focusing on the musical structure and inhibiting potential distractors. This ability involves executive functions (EF), including inhibition control, selective attention, and set-shifting. To examine the relation between EF and beat synchronization, we investigated the impact of beat synchronization training on rhythmic and executive functioning in children with Attention-Deficit Hyperactivity Disorder (ADHD). In a fully remote longitudinal proof-of-concept protocol, 30 children (7-13 years) were randomly assigned to either a finger-tapping rhythmic game (Rhythm Workers, tablet application) or a control game (Frozen Bubble) which had similar auditory-motor demands, but no beat-synchronization (active control condition). The children played the game at home for 300 minutes over 2 weeks. We assessed whether the games are comparable and collected data before and after training on a range of rhythmic abilities using the Battery for the Assessment of Auditory Sensorimotor and Timing Abilities (BAASTA) and EF tasks for inhibition control, selective attention, working memory, and cognitive flexibility. Analyses were conducted as a function of group and time of testing. Preliminary findings suggest that children evaluate the games as similarly difficult and pleasant ($p > .38$). Notably, children who played Rhythm Workers showed improved rhythmic abilities after training, as revealed by enhanced synchronization to the beat of music and a metronome ($p < .05$), relative to children playing the control game. Children who played the rhythmic game had a quicker response on incongruent trials of the selective attention test compared to the control group ($p = 0.02$). These preliminary findings show that training with a rhythm-based serious game can serve to improve rhythmic abilities in children with ADHD and also improve extra-musical abilities such as inhibition control and selective attention, as opposed to a non-rhythmic game.

Author e-mail: hugo.laflamme.1@umontreal.ca

[Back to Schedule](#)

10:15 Neural correlates of multimodal speech perception in cochlear-implanted and normal-hearing adults

Hannah Shatzer*

Toronto Metropolitan University

The addition of visual speech cues generally improves perception of an auditory speech signal, particularly in individuals with hearing loss who may use hearing instruments (e.g., hearing aids, cochlear implants). The neural mechanisms underlying this visual benefit, including the effects of hearing loss and cochlear implantation, are currently being studied. The aim of the current research was to identify common electrophysiological signatures of auditory-only and audiovisual speech perception in deaf and hearing adults, and to investigate potential relationships between those neural patterns and behavioral speech perception abilities across individuals. Postlingually deafened adult cochlear implant users and age-matched normal-hearing control participants completed a word recognition task in auditory-only, visual-only, and audiovisual conditions during electroencephalography (EEG) recording. Auditory-evoked potentials suggested that both groups experienced suppression of early auditory cortex activity with the addition of visual cues. Cochlear implant users who had better recognition of auditory-only speech showed early auditory activity that was more similar to normal-hearing listeners relative to cochlear implant users with poorer speech recognition. Results also indicated that individual lipreading ability and visual benefit correlated with auditory-evoked potential peak amplitudes for audiovisual speech across both normal-hearing and cochlear-implanted participants, suggesting that individuals with stronger visual speech recognition experience more suppression of early auditory activity with visual cues regardless of hearing status. These results indicate that visual speech cues play a critical role in supporting speech perception, and that neural indicators of visual benefit are present in both normal-hearing and cochlear-implanted listeners.

Author e-mail: hannah.shatzer@ryerson.ca

[Back to Schedule](#)

10:30 Learning mechanisms in phonetic cue weighting: What we learn from a single cue

Adam Bramlett

Carnegie Mellon University

Kaori Idemaru

University of Oregon

Vsevolod Kapatsinski*

University of Oregon

Cue reweighting refers to adjusting associations of perceptual cues like voice onset time (VOT) and fundamental frequency (F0) based on perceptual experience. For example, to adapt to Korean-accented English, native English listeners need to increase reliance on F0 and reduce reliance on VOT. We compare the predictions of two computational models of learning for this process, the most influential model of error-driven learning (RW; Rescorla & Wagner, 1972) and a recent model we have developed based on prior cue reweighting results (Harmon, Idemaru & Kapatsinski, 2019, Cognition). The models differ in two ways. First, RW reassociates specific cues (like VOT=45 ms) with unexpected outcomes, whereas HIK reweights perceptual dimensions like VOT and F0. Second, RW downweights a predictive cue when it becomes unpredictable, whereas HIK downweights a predictive dimension only if there is a more predictive dimension to rely on.

We presented native English participants (N=240) with a single VOT value paired with 50/50 beer/pier feedback with a pretest/posttest design. F0 was either predictive or not.

Participants showed learning only when trained on a priori informative VOT values (5 ms or 45 ms but not 25 ms). This training weakened the association between the experienced value of VOT and the a priori expected response (/b/ for VOT=5, or /p/ for VOT=45). The associations generalized to similar values of VOT (15 and 35 respectively). But VOT as a dimension was not downweighted: learning that VOT=45 unexpectedly often cues /b/ did not lead participants to infer that VOT=5 will unexpectedly often cue /p/. These results are consistent with RW. However, participants show more learning about VOT when F0 is predictive of feedback, i.e., when they have another cue to rely on during training. This is consistent with HIK. We outline ways to combine these models to yield the observed pattern of results.

Author e-mail: vkapatsi@uoregon.edu

[Back to Schedule](#)

10:45 Distributional learning of a novel speech contrast by bilingual and monolingual listeners

Siyu Lin*

Grant M. Berry

Joseph C. Toscano

Emory University

Villanova University

Villanova University

Bilingual speakers have exposure to a wider range of phonetic distinctions than monolingual speakers of the same languages. As a result, they may have more experience perceiving fine-grained acoustic differences and tracking distributional statistics of speech sound categories. The broader experience bilinguals possess may thus facilitate the acquisition of new category boundaries when exposed to novel phonemic contrasts. The current study focuses on differences between highly-proficient Spanish-English bilinguals and monolingual English speakers in acquiring a novel phonetic distinction that is not native to either language: the French /y/-/i/ distinction. This distinction depends on rounding of front vowels, cued by variation in the third formant (F3). To assess listeners' ability to learn this distinction, 20 monolingual and 20 bilingual participants completed a distributional learning experiment. Stimuli consisted of synthetic speech sounds varying in F3 between /y/ and /i/, and a bimodal distribution of sounds along the F3 continuum was created with peaks corresponding to prototypical /y/ and /i/ sounds. During training, listeners were asked to indicate whether sounds drawn from this distribution were /y/ or /i/. Before and after training, listeners completed a flanked four-interval forced choice (4IAX) discrimination task. We predicted that the bilingual listeners would perform better than the monolingual listeners in post-training discrimination performance and in acquiring the novel /y/ category (based on their /y/-/i/ categorization responses during training). Although there were no significant differences between the two groups in the discrimination task, the categorization results revealed an interaction between F3 step and language group, indicating that bilingual listeners had sharper categorization functions. This suggests that they acquired the novel category better than monolingual listeners. This finding supports the idea that additional linguistic experience can lead to more efficient tracking of distributional information in speech, which benefits speakers in learning novel phonetic categories.

Author e-mail: siyu.lin@emory.edu

[Back to Schedule](#)

11:00 **The role of rhythm complexity in cardiac dynamics during auditory perception and production**

Shannon E. Wright* McGill University

Selva Asgharizadeh McGill University

Caroline Palmer McGill University

Perception and production of auditory rhythms are key capacities for human behaviours such as speech and music. Auditory rhythms range in complexity of tone durations; complex rhythms (based on non-integer duration ratios) are more difficult to perceive and to perform than simple rhythms (integer ratios). Differences in neurophysiological (EEG) activity have been observed for musicians' perception and production of simple and complex rhythms (Mathias et al, 2020). Music perception and performance also modulate cardiac and respiratory activity. This study investigates how rhythm complexity affects cardiac dynamics during auditory perception and production.

Twenty-one adults with musical training (mean years instruction = 13.03) listened to and performed simple and complex auditory rhythms while their cardiac activity was recorded. The simple rhythm was composed of a high tone and a low tone whose temporal intervals formed a 1:2 ratio and the complex rhythm was composed of the same tones whose temporal intervals formed a 3:2 ratio. Cardiac activity was measured during perception and production tasks. Participants listened to the auditory rhythms and indicated whether there was a missing tone. Then participants performed the low tones of the auditory rhythms by synchronising their taps with the high tones on a midi keyboard. Listeners were equally accurate at identifying the missing tone in simple and complex rhythms. Tapping synchronisation was less accurate with complex rhythms than with simple rhythms. Mean heart rate was slower during rhythm perception than during production, for both simple and complex rhythms. Heart rate variability showed an interaction between task and rhythm type; variability was lowest during complex rhythm production. Similar to changes in mean heart rate, recurrence quantification analysis showed patterns of more recurrent cardiac activity during rhythm production than perception. These results suggest that cardiac activity is modulated by both task (perception, production) and by rhythm complexity.

Author e-mail: shannon.eilyce.wright@mail.mcgill.ca

[Back to Schedule](#)

11:15 Undetectable Very-Low Frequency Sound Increases Dancing at a Live Concert

Michael Hove*	Fitchburg State University
Daniel Cameron	McMaster University
Dobromir Dotov	McMaster University
Erica Flaten	McMaster University
Daniel Bosnyak	McMaster University
Laurel Trainor	McMaster University

In music, low frequency sound is associated with the musical quality that makes us want to dance—groove. Bass-heavy music tends to be rated high in groove, and bass instruments typically provide the pulse to which dancing synchronizes. Low frequencies confer advantages in perception and movement timing, and elicit stronger neural responses compared to higher frequencies, suggesting that the mechanism underlying these low-frequency effects involves superior auditory-motor communication. Notably, low frequency sound is also processed via vibrotactile and vestibular pathways and stimulation of these modalities can modulate musical rhythm perception. Anecdotal accounts describe intense physical and psychological effects of low frequencies, especially in electronic dance music. However, we do not know if these associations extend to direct causal effects of low frequencies in complex, real-world, social contexts like dancing at concerts, or if low frequencies that are not consciously detectable can affect behavior. Here, we used specialized speakers to manipulate the presence of very low frequencies (VLF; 8-37 Hz) in an ecologically valid electronic music performance and measured audience members' body movement with motion capture. The VLF levels were below or near thresholds for auditory perception, and those that were above-threshold were likely inaudible due to masking by the non-VLF sound in the music. Audience members moved more when VLFs were ON, suggesting an unconscious effect on behavior, possibly via vestibular and tactile processing.

Author e-mail: michaeljhove@gmail.com

[Back to Schedule](#)

New Models of Human Hearing via Machine Learning

Josh McDermott*

MIT

Humans derive an enormous amount of information about the world from sound. This talk will describe our recent efforts to leverage contemporary machine learning to build models of our auditory abilities and their instantiation in the brain. Such models have enabled a qualitative step forward in our ability to account for real-world auditory behavior. But they also exhibit substantial discrepancies with human perceptual systems that we are currently trying to understand and eliminate.

Author e-mail: jhm@mit.edu

[Back to Schedule](#)

2:45 Acoustic change detection during associative auditory learning and explicit memory is related to enhanced signal detection at retrieval

Manda Fischer*	University of Toronto; Rotman Research Institute
Morris Moscovitch	University of Toronto; Rotman Research Institute
Claude Alain	University of Toronto; Rotman Research Institute

Associative memory between a soundscape and an embedded lateralized pure tone target facilitates target detection at retrieval. We tested the hypothesis that benefits in long-term associative memory on performance are related to 1) the salience of the target during learning and 2) explicit memory for the soundscape.

First, we created associations between soundscapes (e.g., waterfall sound) and the location (left/right) of an embedded pure tone target. To ensure that incidental learning of the associations could proceed, participants made irrelevant judgements about the soundscape or target tone. We did not make any reference to the location of the targets or their relation to the soundscapes. At retrieval, participants detected the target as quickly as possible and answered questions about their memory for the soundscapes, target, and associations. To track learning-related changes, we measured neuroelectric brain activity at learning and focused on the acoustic change complex (ACC), a component that reflects automatic processing of the target tone at the level of the auditory cortex.

Detection accuracy at test was associated with greater ACC amplitudes at learning. Further, target detection at test varied as a function of explicit memory for the soundscape.

Memory for the soundscape may facilitate processing efficiency at test. Larger ACC amplitudes may reflect greater sensitivity to acoustic changes at the level of the auditory cortex, which may in turn facilitate associative memory encoding.

This study examined the contributions of automatic processing of the target at encoding and memory for the soundscape on target detection at retrieval. Together, larger ACC amplitudes may reflect enhanced target saliency and enable associative memory at encoding, while explicit memory for the soundscape may support general processing efficiency to aid target detection at retrieval.

Author e-mail: manda.fischer@mail.utoronto.ca

[Back to Schedule](#)

**3:00 Perceived Similarity and Recognition Memory for a Set of Novel
 Auditory Stimuli**

Nathan Gillespie*
Gregory Cox

University at Albany, SUNY
University at Albany, SUNY

Research in categorization and recognition memory typically employs isolated, static stimuli, whereas many events that we experience in life (e.g., musical or linguistic events) are extended in time and potentially overlapping. Studying these more ecologically valid events is challenging because their complexity makes it difficult to relate their physical dimensions with their psychological representations. We began to address this gap by developing a set of novel auditory stimuli with experimentally controlled physical features, such that differences among them were analogous to differences in timbre. To accomplish this, we generated the stimuli by manipulating the frequency bands above a 200hz fundamental of eight electronically generated sounds so that they exhibited varying degrees of spectral overlap. Across two studies, participants rated similarity between pairs of these sounds and engaged in a recognition memory task. We applied non-metric multidimensional scaling to similarity ratings to obtain a three-dimensional psychological representation of these stimuli. The first dimension appeared to correspond to timbral roughness and the second to timbral brightness, while the third did not admit a simple verbal label. There were also individual differences in the degree to which participants attended to each of these dimensions, potentially as a function of musical expertise. The representation inferred from similarity ratings predicted recognition memory performance for single probe sounds following sequential presentation of two sounds, consistent with similarity-based exemplar models of memory. We discuss implications for combining these stimuli in sequential or simultaneous patterns, treating them as an “alphabet” to construct sounds that mimic the complexity of realistic auditory events while retaining well-established and controllable physical and psychological dimensions.

Author e-mail: ngillespie@albany.edu

[Back to Schedule](#)

3:15 **Is Auditory Preemption a Thing? Five Reasons to be Skeptical**
Michael A. Nees* Lafayette College

Auditory preemption applies to the case where a person is engaged in an ongoing visual task and intermittently is required to engage in an additional discrete task. Some research has suggested that an interruption from an auditory discrete task is more damaging to the ongoing visual task than an equivalent interruption from a visual discrete task—a phenomenon that has been described formally as Auditory Preemption Theory. In this critical comment, I review the definition and history of the notion of auditory preemption and survey the literature that has purported to show auditory preemption. I present five reasons to be skeptical about auditory preemption: (1) The original data presented to introduce the term were not analyzed formally and did not consistently support auditory preemption; (2) One of the primary pieces of evidence cited to support auditory preemption relies on interpreting nonsignificant results as if they were significant; (3) The most cited studies in support of auditory preemption have not produced conclusive evidence of auditory preemption; (4) A meta-analysis found no evidence to support the primary prediction of auditory preemption theory; and (5) Auditory preemption has shortcomings with respect to fundamental theoretical functions, including explanation, prediction, and falsifiability. I argue that the notion of auditory preemption should be abandoned unless conclusive evidence of auditory preemption emerges.

Author e-mail: neesm@lafayette.edu

[Back to Schedule](#)

**3:30 Triggering misophonia: The importance of spectral information,
temporal information, and action identification**

Savvas Kazazis*

McGill University

Iza Korsmit

McGill University

Stephen McAdams

McGill University

Misophonia is characterized by severe negative emotional responses to specific environmental sounds. In this experiment, we investigate the importance of spectral and temporal acoustic information, as well as the role of action identification (e.g., chewing) in triggering Misophonia. Eighteen participants with severe misophonia completed the experiment (mean score of 44.83 out of 63 on the Misophonia Assessment Questionnaire). Interestingly, 50% of our participants also experience ASMR. Our stimuli consisted of recorded sounds that were either common misophonic triggers or neutral sounds (that are not expected to trigger misophonia). Both types of stimuli were presented in three formats; unmodified (as recorded), spectrally modified, and temporally modified. Spectral modifications were achieved through spectral whitening while preserving the global amplitude envelope. Temporal modifications were achieved by scrambling short audio segments which allowed us to preserve the short-term spectral structure but distorted the global amplitude envelope. Participants rated how triggered they were by each sound (i.e., averseness), and were asked to identify the action category of each sound. The unmodified trigger sounds were rated to be more aversive than neutral sounds ($p < 0.0001$). Linear mixed modelling to predict averseness of misophonic triggers revealed that the main effects of modification type and identification were significant ($p < 0.0001$). For modification type, post-hoc tests showed that the differences between the averseness ratings of the unmodified and temporally modified sounds were not significant ($p > 0.29$). However, the ratings of the spectrally modified sounds were significantly lower than the ratings of both the temporally modified sounds ($p < 0.005$) and the unmodified sounds ($p < 0.001$). Finally, regarding identification, the stimuli that were incorrectly identified were on average rated as less aversive than the ones that were correctly identified ($p < 0.0001$). This shows that both identification and spectral information play an important role in triggering misophonia.

Author e-mail: savvas.kazazis@mail.mcgill.ca

[Back to Schedule](#)

4:00 The Auditory Looming Bias is Influenced by Visual Information

Maggie K. McCracken*

University of Utah

John G. Neuhoff

College of Wooster

Looming objects generate both auditory and visual information that is perceptually important to an observer. However, these sensory domains are not perceived equally. Visual arrival time estimations are relatively accurate, while estimations based on audition are underestimated. This auditory looming bias is also studied through loudness change, as looming sounds are perceived as changing more in the loudness than equivalent receding sounds. Previous research examining multisensory integration of looming objects typically uses audiovisual estimations. Here, we are interested in how visual motion influences auditory looming perception. In two experiments, participants judged loudness change for looming and receding sounds that were presented with coincident visual motion. When presented with a greater visual change, sounds were perceived as having a greater loudness change. The results of the first experiment suggest that auditory perception is susceptible to changes in the visual domain. To further test the relationship between the two sensory domains, we presented participants with looming and receding sounds, half of which were presented with coincident visual motion. Sounds were perceived as changing more in loudness when presented unimodally. Furthermore, participants exhibited a larger auditory looming bias in the unimodal condition, as shown through the difference between looming and receding estimations. This second experiment implies that motion estimations are most cautious when presented only with the more ambiguous sensory domain. Taken together, this study shows that the auditory looming bias is influenced by vision and provides additional evidence supporting an adaptive bias in the integration of looming objects.

Author e-mail: maggie.mccracken@psych.utah.edu

[Back to Schedule](#)

4:15 A cue or a distractor? Automatic attention in spatial discrimination

Rene Sebena

P. J. Safarik University, Kosice, 04001, Slovakia

Norbert Kopco*

P. J. Safarik University, Kosice, 04001, Slovakia

A behavioral experiment and EEG recordings were combined to examine how directing automatic auditory spatial attention affects performance in a location discrimination task while the eyes fixate a neutral location, and whether the effect is cue modality dependent. Analyses on behavioral data found 1) better performance with auditory valid vs. invalid cues, 2) no difference in performance with visual valid vs. invalid cues; and 3) bias toward perceiving the sounds as moving away from the fixation direction, the strength of which depended on cue modality and validity. EEG analyses revealed that the late components of the target-elicited ERPs covaried with the behavioral performance. The analysis of cue-elicited ERPs examined the auditory-evoked occipital response contralateral to an auditory cue (ACOP), previously reported as a correlate of attentional processing. While analysis with respect to head-centered reference frame of the interval 300-500 ms did not find clear evidence of ACOP in the eye-centered reference frame, we observed a relatively weak trend for contralateral positivity in the interval 300 – 400 ms. These results demonstrate that the both eye fixation and auditory cue location affect auditory spatial discrimination, while the visual cue has only a small effect. However, the auditory cue effect is more a distracting one when the cue is invalid than an enhancing one when the cue is valid.

Author e-mail: nkopco@gmail.com

[Back to Schedule](#)

4:30 Inferring object interactions from sound

Vinayak Agarwal*

MIT

Josh McDermott

MIT

Object interactions – collisions, scraping and rolling – create many of the sounds that we hear in the world around us. These sounds are generated via lawful physical dynamics. Anecdotally, humans possess some intuitive knowledge of the physical generative processes underlying sound production, but little is known about the extent and nature of this knowledge. We developed new methods to synthesize sounds of impacts, scraping and rolling from physical variables, and used them to study the human ability to make inferences about physical properties from contact sounds. The results indicate that humans estimate physical variables such as mass and elasticity, and are sensitive to the physical consistency of different segments of sound that are heard over time. The results are suggestive of auditory intuitive physics, in which humans infer physical variables from sound and store them over time in object representations.

Author e-mail: vinayaka@mit.edu

[Back to Schedule](#)

Poster Abstracts

Auditory Memory and Attention (1-4)

1 ERP correlates of auditory peak shifts in stimulus generalization

Chelsea N. Joyner*

Kansas State University

Matthew G. Wisniewski

Kansas State University

The peak shift effect is a widely observed phenomenon in which discrimination learning causes generalization gradients for learned responses to peak at a novel stimulus rather than a trained one. Associative learning and non-associative representational plasticity have been used as explanations of the peak shift. Associative theories posit that the peak shift arises from reweighting connections between stimulus representations and decision/response outputs. Non-associative representational plasticity perspectives explain the peak shift as resulting from changes in stimulus representations themselves. We investigated behavioral and event-related potential (ERP) expressions of the peak shift effect, the latter being used to characterize the processing stage(s) in which peak shifts arise. Stimuli consisted of 7 different frequency modulation (FM) rates ranging from “slow” (4.14, 6.9, 7.2 octaves/s) to “fast” (9.52, 9.94, 16.56 octaves/s). Subjects were trained with feedback to respond “Target” for a 8.28 octaves/s FM rate and “Non-Target” to 6.9 octaves/s. Test trials using FM rates from the entire stimulus continuum were interspersed among these training trials. Analyses of “Target” responses on test trials revealed a gradual increase in peak shift as discrimination learning progressed. This was evident in more “Target” responses to a shifted 9.52 and 9.94 octaves/s FM rate than the trained target rate. Interestingly, two components of the ERP response to test stimuli appeared to parallel these changes seen in behavior: the P2 and a Late Posterior Positivity (LPP). Data is discussed in regard to what these different components may mean for the processes involved in perceptual discrimination learning, and potential contributions from multiple learning processes in generating peak shifts.

Author e-mail: cjoyner@ksu.edu

[Back to Poster Session](#)

2**Accuracy and Precision of Reproduced Short Term Memories**

Michael A. Tollefsrud*

Kansas State University

Matthew G. Wisniewski

Kansas State University

The most prevalent conceptualizations of working memory capacity see it as divided into several fixed capacity subspaces (slot models) or as a flexible resource that can be divided among items and features (resource models). To distinguish these two possibilities, recent auditory working memory studies have employed method of adjustment paradigms in which listeners reproduce sounds in real-time (e.g., by adjusting frequency) to match their memory of previously presented items. Several have reported that the precision with which subjects reproduce a sound after a delay decreases with increasing memory load, suggesting that capacity can be divided flexibly among items. We use a similar approach to examine whether capacity can be divided flexibly between two acoustic dimensions of a single item. In two experiments, listeners were presented with sounds having a randomly selected amplitude modulation (AM) rate (4 - 8 Hz) and carrier frequency (500 Hz - 1000 Hz). They were precued on each trial to pay attention to AM, carrier frequency, or both dimensions. After a retention interval, they attempted to reproduce the sound by moving their finger on an XY MIDI pad that changed these dimensions of a novel sound in real-time. Experiment 1 revealed no reduction in precision or accuracy of these adjustments when subjects were cued to both dimensions compared to when they were cued to a single dimension. Experiment 2 replicated the methods of Experiment 1, but also manipulated the number of items to be maintained (1 vs. 3). Though memory load significantly decreased the precision and accuracy of listeners' reproductions, there was still no impact of the number of cued dimensions. Though the results support the position that capacity can be allocated flexibly between items, no evidence of flexible allocation within items was obtained.

Author e-mail: matollef@ksu.edu

[Back to Poster Session](#)

3 Exploring the Disruptive Role of Background Speech on Short-term Memory: What Role for Whispering?

Kathleen C. McCulloch*	University of Central Lancashire, UK
Tom A. Campbell	Tampere University, Finland
Zoe Hughes	University of Central Lancashire, UK
John E. Marsh	University of Central Lancashire, UK

The mere presence of background sound disrupts visual-verbal serial short-term memory: the irrelevant sound effect. Evidence suggests that the acoustic properties of background sound such as changes in pitch and timbre are responsible for this disruption, but other acoustic factors such as intensity play little, if any, role. However quieter sounds presented in a whispered, as compared to normally phonated, voice may produce greater disruption due to their intrigue and/or the tendency for whispered speech to elicit an auditory sensory meridian response (ASMR). In the current study participants undertook a serial recall task while presented with whispered or normally phonated, isolated meaningful sentences. A control, quiet condition was also deployed. Between blocks of trials within the study, participants also indicated whether they had experienced an ASMR during exposure to sound and were asked to determine which sound (normally phonated or whispered speech) triggered it. The results demonstrated that whispered sentences were more disruptive than normally-phonated sentences, suggesting that they produced greater attentional diversion. Further, participants self-reported whispered sentences to produce ASMR with a higher frequency than normally-phonated sentences. However, the tendency to experience an ASMR was not associated with the magnitude of distraction from whispered against normally-phonated sentences. Implications of the findings for dynamic in-vivo social settings (e.g., the workplace or classroom) are discussed.

Author e-mail: KMcculloch@uclan.ac.uk

[Back to Poster Session](#)

4 Cognitive control of auditory attentional capture: Examining the interaction of contextual factors and trait capacity for steadfast task-engagement

John E. Marsh*	University of Central Lancashire, Preston, UK
Zoe D. Hughes	University of Central Lancashire, Preston, UK
Federica Degno	Bournemouth University, Bournemouth, UK
Philipp Ruhnau	University of Central Lancashire, Preston, UK
Robert W. Hughes	Royal Holloway, University of London, Egham, UK

The detection of a task-irrelevant auditory deviant typically results in a switch of attention to the deviant. However, response to contextual factors (e.g., high focal-task difficulty) and a stable disposition for strong attentional control—as measured by working memory capacity tasks—is thought to temper this disruption by facilitating the active blocking of the switch of attention to the deviant following its detection. The present research examines how contextual factors and trait capacity for attentional control interact in the cognitive control of auditory attentional capture. First, we examine whether a proactive, as compared to a reactive, focal-task encoding mode—manipulated by varying the predictability of the focal-task stimuli—reduces the disruptive effect of an auditory deviant. Second, we investigate whether any attenuating effect of a proactive encoding mode on attentional capture by a deviant is modulated by individual differences in working memory capacity. The study sheds light on the potentially interacting roles of contextual factors and trait capacities in mitigating auditory attentional capture.

Author e-mail: JEMarsh@uclan.ac.uk

[Back to Poster Session](#)

5 **Dimensionality of Natural Auditory Scene Perception: A Factor Analysis Study**

Margaret A. McMullin*
Nathan C. Higgins
Brian Gygi
Joel S. Snyder

University of Nevada, Las Vegas
University of South Florida
East Bay Institute for Research and Education
University of Nevada, Las Vegas

Theories of auditory and visual scene analysis suggest the perception of a scene relies on the identification and segregation of objects within it, resembling a detail-oriented processing style. However, it is possible that a more global-oriented process may occur in parallel when we evaluate auditory scenes. There is evidence in the visual domain for global properties that enable scenes to be rapidly recognized, even without recognizing the individual objects comprising the scene. It is our understanding that a similar line of research has not been explored in the auditory domain. Therefore, we evaluated the contributions of high-level global and low-level acoustic information to auditory scene perception. A secondary aim is to increase the field's ecological validity by utilizing our collection of high-quality auditory scenes. Participants rated scenes on 8 global properties (e.g., open vs. enclosed, natural vs. human-influenced, etc.) and an acoustic analysis evaluated which low-level features predicted the ratings. We submitted the acoustic measures and average ratings of each scale to separate exploratory factor analyses, which explained 57% and 64% of the variance in the data, respectively. Eight regressions calculated which acoustic features predicted global property ratings; each scale was predicted by at least one acoustic variable ($R^2 = 0.33-0.87$). These results provide preliminary evidence for the ability to perceive auditory scenes from a global perspective. Some of the acoustic measures predicted performance on the rating task, suggesting global variables may be processed at a high level where acoustic features are abstracted out. Alternatively, scene representations may be transformed through many stages of processing in the auditory ventral stream, similar to what has been proposed to occur in the ventral visual stream. The results of this study and the open availability of our scene collection will make future studies on perception, attention, and memory for natural auditory scenes possible.

Author e-mail: mcmulm1@unlv.nevada.edu

[Back to Poster Session](#)

6

Auditory Judgments of Egocentric and Allocentric Motion and Motion Direction

Michael K. Russell*

Bellevue University

We are regularly exposed to innumerable sound-producing objects and the spatial relationship between observer and sound source regularly undergoes change. In certain instances, a sound source will change location while an observer remains stationary. In other instances, it is the observer that changes location while the sound source remains stationary. The findings of previous studies suggest individuals are capable of accurately determining the movement of an unseen, moving sound source. Previous research has also discovered that a number of acoustic factors (e.g., intensity, frequency, ratio of direct-to-indirect sound) impact the perception of movement. Although the sound contacting the ear may be informative about the change in the observer-source spatial relationship, reasons exist for believing that the acoustic energy array does not contain information specifying the origin of the change. Theoretically, individuals should be incapable of discriminating between egocentric events (i.e., events that reflect the observer as the agent of change) and allocentric events (i.e., events that reflect the sound-producing object as the agent of change). It was hypothesized that participants in the present study would be largely incapable of differentiating between egocentric and allocentric events yet highly capable of differentiating between approach and withdrawal events. Participants in the present study listened to 40 brief audio events. Following each event, the participants were asked to identify the agent of change (sound source or observer) and the direction of motion (approach or withdrawal). Discussion will be given to the extent to which acoustic information alone is capable of specifying the origin of change in the observer-source spatial relationship. The findings of the present study have implications on our understanding of how we perceive the world by sound.

Author e-mail: mirussell@bellevue.edu

[Back to Poster Session](#)

Despite theoretical assumptions and empirical approaches, empty space should be considered a limiting case. The simple fact is that the settings we perceive and the world we act within is nearly always cluttered. The presence of clutter can dramatically alter the sound reaching the ear and influence auditory judgments. Despite the common existence of clutter in real-world settings, relatively few investigations have examined the impact of clutter on auditory perceptions and aurally guided actions. Recently, Russell and Brown (2019) discovered that participants are generally capable of detecting the presence of an obstruct. When the amount of obstruction exceeded 50%, participants believed an object was between them and an unseen sound source. When the amount of obstruction was less than 50%, participants believed the sound source was unoccluded. It is worth noting that study was conducted in-person and participants obtained visual information relating to the setting, sound source location, obstruct composition and size, and more. Since an individual's visual experience has been shown to affect auditory perception, it is possible judgments of occlusion were influenced by the visual information picked up by the participants. In the present study, participants were exposed to audio recordings of an unobstructed, partially obstructed, or fully obstructed sound source. The task of the participants was to simply report whether the sound source was occluded. An increase in actual occlusion was expected to yield an increase in reports of occlusion. It was further expected that the absence of visual information would significantly impair the detection of an obstruct particularly when the sound source was partially occluded. The findings of the present study will be discussed in terms of the ability of individuals to detect occlusion, the impact of acoustic factors on occlusion perception, and the significance of visual information on auditory spatial perception.

Author e-mail: mirussell@bellevue.edu

[Back to Poster Session](#)

8 Streaming Sound Texture in Auditory Scenes

Jarrod M Hicks*

MIT

Josh H McDermott

MIT

Sound textures are created by the superposition of many similar acoustic events (e.g., rain falling, birds chirping, or people clapping) and are thought to be represented in the auditory system by statistics that summarize acoustic information over time. Real-world auditory scenes frequently contain multiple concurrent textures (as when birds chatter next to a babbling brook), raising the question of whether listeners can “hear out” (i.e., stream) individual textures. We sought to characterize “texture streaming” by asking whether listeners can estimate the number of sound texture sources in an auditory scene. In the first experiment, participants heard auditory scenes composed of one or two real-world textures and judged the number of distinct sound sources. Listeners performed above chance, tending to correctly judge the number of sources in each auditory scene. Inspection of judgments for individual scenes revealed consistent patterns of errors, indicating that particular combinations of textures tended to be mistakenly heard as a single stream. In a follow-up experiment, we asked participants to rate the similarity of texture pairs from the streaming experiment and found similarity ratings to be partially predictive of streaming judgments. However, a substantial portion of the explainable variance in streaming judgments could not be predicted from perceptual similarity of the source textures, suggesting additional (as yet not understood) principles of perceptual organization. Together, these experiments demonstrate the phenomenon of texture streaming—a neglected aspect of auditory scene analysis in which listeners are able to stream concurrent textures in auditory scenes.

Author e-mail: jmhicks@mit.edu

[Back to Poster Session](#)

9 Predictive coding : Electrophysiological signature of timing predictions

Amour Simal*	University of Montreal; International Laboratory for Brain, Music, and Sound Research (BRAMS)
Robert Zatorre	Montreal Neurological Institute; McGill University; BRAMS
Pierre Jolicoeur	University of Montreal; BRAMS; CRIUGM

In the auditory domain stimuli unfold quickly, so rapid pre-attentive coding of predictions could help to disambiguate contextual information and to orient attention efficiently. Our study builds on a previous study by Simal et al. (2021). They showed larger N1 and P2 event-related potential components for tones predicting the presentation of additional subsequent tones. Our goal here was to test if a similar effect exists for predictions of “when” the next stimuli will be presented. We used three types of five-tone sequences with distinct temporal regularities. They were designed so the second tone carried information, and allowed recognition of the regularity, and thus, prediction. Tones lasted 200 ms and the ISI could be 200 ms (short), 300 ms (medium), or 400 ms (long). Intervals could be isochronous (all 300 ms), short-long-short-long, or long-short-long-short. Twenty-two participants heard two sequences of 5 tones, identical in timing, interspersed by a silent interval (2 s). They decided whether the tone pitches in the two sequences were the same or different. In Experiment 1, the three types of regularities were randomly intermixed. When the second tone carried information, in the first sequence, we observed a frontocentral positivity from tone onset to the onset of the N1. We confirmed this was not caused by low-level effects, nor by memory processes by replicating the experiment with blocked conditions, making local prediction superfluous. Because no effects were seen on the N1 or the P2, we believe the mechanisms in play to keep track of the timing in sound sequences differ from those used to evaluate sequence length. Moreover, the temporal regularities were not relevant to the task and overlooked by most participants. The results suggest predictions on the timing of auditory events are rapid, pre-attentive, and automatic.

Author e-mail: amour.simal@umontreal.ca

[Back to Poster Session](#)

10 Pleasantness altered by misidentification of everyday sounds

Laurie M. Heller*	Carnegie Mellon
Jessica Smith	Carnegie Mellon
Urszula Oszczapinska	Carnegie Mellon

Three studies investigated the influence of source identification on the emotional response to everyday sounds. In all three studies, the sounds spanned a variety of emotional categories, such as Pleasant, Neutral, Unpleasant, and Misophonic. Misophonic sounds are a subset of sounds that are especially unpleasant to people who suffer from misophonia. The first two studies, from Heller & Smith (2022), paired sounds from different emotional categories based on similar causal properties with the goal of inducing misidentifications. These studies obtained sound identification accuracy, ratings of sound pleasantness, and causal properties, such as actions and materials. The first study found that a sound's pleasantness would increase or decrease accordingly when it was misidentified as being from a different emotional category. The second study replicated the results of the first one while inducing more misidentifications by imposing spectral degradation in the form of temporal envelope vocoding with a noise carrier. A third study extends these findings to a broader set of environmental sounds (ESC-50, Piczak 2015). [Funded by REAM foundation.]

Author e-mail: hellerl@andrew.cmu.edu

[Back to Poster Session](#)

11 **A preliminary evaluation of potential perceptual contributions from the skull**

Laura Reinert*

Michael D. Hall

Michael S. Gordon

James Madison University

James Madison University

William Paterson University

Amplification from the pinna, auditory canal, and ossicles is assumed to be sufficiently large to render potential perceptual contributions from bone conduction (BC) through the skull negligible (e.g., Stenfelt et al., 2002). However, the observance of large differences in BC filtering as a function of frequency (with peak amplitudes correlating with BC thresholds; see Gordon et al., 2019) raises the possibility that BC signals could be perceived when mixed with air-conducted (AC) signals (through the ossicular chain). The current investigation took initial steps toward addressing this possibility.

Third-octave band filter settings were created in *MaxforLive* to separately match attenuation levels from 27 skulls that were measured by Gordon, et al. (2019). A sawtooth wave ($F_0=43.06$ Hz) was submitted to the filter bank to simulate BC signals. These were subsequently mixed with the original sawtooth, such that the relative peak amplitude of BC signals was -8, -20, -32, or -44 dB. Mixed signals were equated for average RMS amplitude. Stimuli were presented in a 4IAX discrimination task. Participants were asked if the first or second tone pair on each trial contained an altered/BC signal; the original sawtooth constituted the remaining three tones on a trial.

At the time of abstract submission, analyses from the -8 dB and -20 dB conditions were complete. Sensitivity (d') remained high (mean=3.39) across all filter settings at -8 dB, and only dropped to unreliable (< 1) levels for one skull at -20 dB. This pattern will be compared with data from the remaining attenuation conditions. Should discrimination performance be maintained when BC signals approach intensities that reflect their estimated amplitudes relative to AC signals, then this would suggest audible contributions from the skull under normal (field-based) listening conditions, and particularly when the AC signal is reduced (such as when ears are covered). Implications will be discussed.

Author e-mail: reinerlm@jmu.edu

[Back to Poster Session](#)

12 **Categorization of environmental sounds according to the manner of
contact between solids**

Michael D. Hall

Benjamin H. Stone*

Michael K. Russell

James Madison University

James Madison University

Bellevue University

Listeners frequently access invariant stimulus information to identify critical aspects of environmental sounds. For example, listeners categorize events well above chance levels based upon the contact between states of matter that produced them (e.g., solid-liquid; Russell & Siebenaler, 2021). The current investigation extends these findings by evaluating the extent to which adults use structural properties within acoustic signals, including changes in amplitude, to specify the nature of solid-solid contact. Seven unique events were digitized for each of seven forms of contact— bouncing, breaking, rolling, scraping, slashing, sliding, and striking. Events were matched for average amplitude and randomly presented. The task of listeners was to identify the form of contact. Categorization (across six trials per event) was quite accurate, with listeners reliably identifying most event categories. However, rolling, scraping, and sliding categories were regularly confused with each other. This might be due in part to amplitude-matching procedures, which artificially raised the intensity of sliding events. Acoustic analyses further revealed distinct amplitude envelopes that were associated with each category. An ongoing experiment seeks to determine the degree to which amplitude envelopes alone might indicate the nature of solid-solid contact. Envelopes were sampled (to the msec) from the event recordings and were imposed upon brown noise. Listeners categorized 10 randomized instances of each original sound and each noise sample across counterbalanced blocks of trials. Certain categories of noise samples (e.g., striking) are expected to be accurately labelled, indicating that their amplitude envelopes could be sufficient for classifying particular forms of contact. Conversely, many samples (e.g., sliding) should be inaccurately categorized, indicating that amplitude is insufficient to permit reliable classification. Implications for ecological and psychoacoustic approaches to understanding environmental event perception will be discussed.

Author e-mail: stone2bh@dukes.jmu.edu

[Back to Poster Session](#)

13 **Time, intensity, and frequency contributions to clasp perception in cyclic musical rhythms**

Benjamin H. Stone*
Michael D. Hall

James Madison University
James Madison University

Generally, rhythms predictably repeat, establishing a 'pulse'. Perceptual location of pulse is impacted by several factors. These include changes in intensity, relative timing, and frequency. For example, listeners perceive intense tones as longer, and higher tones as more intense (see Povel & Okkerman, 1981). Furthermore, the perceived beginning of a rhythm, or clasp, tends to either follow the largest gap of the cycle or reflect a beat with lower frequency (see Yu, Getz, and Kubovy, 2015). The current investigation extended these results by combining changes across all three parameters to determine their relative contribution to clasp determination in musical rhythms. Cyclic drum patterns were created from a repeated drum hit/beat within a 5/4 measure. Stimulus manipulations were restricted to a single beat within the sequence, including changes to its relative timing (-150, 0, +150 ms), intensity (-12, 0, +12 dB), and/or frequency (-3, 0, +3 semitones, which impacted spectral centroids). Orthogonal combinations of values resulted in 27 rhythmic variations. On each trial, a pattern repeated 8 times, with visual numerals indicating the current beat. Participants then determined the clasp. Each beat had an equal likelihood of appearing first. This minimized potential bias to identify the first beat as the clasp, which only was observed in the absence of stimulus manipulation. Listeners typically reported the clasp as the beat following the largest gap. Remaining manipulations impacted performance only when such a gap was absent, with higher-intensity beats identified as the clasp. If intensity was equivalent across beats, then listeners identified any spectral adjustment as the clasp. These results suggest a hierarchy of cues to perceptual grouping given the fact that all manipulations were perceptually salient and musically relevant. Implications for models of perceptual grouping, as well as for the application of scene analysis cues to rhythmic stimuli, will be discussed.

Author e-mail: stone2bh@dukes.jmu.edu

[Back to Poster Session](#)

14 **Auditory Statistical Learning in Classical and Jazz Musicians**

Erica R. Knowles*

Berklee College of Music

Mateo Larrea

Berklee College of Music

Sheena Odabashian

Berklee College of Music

The ability that music has to give us pleasure and to manipulate our emotions stems from the creation and violation of our expectations of what will come next. Humans acquire the knowledge of the structural rules of music through passive exposure and have been found to be sensitive to these rules. This processing can be further shaped through musical training. Previous research has demonstrated that musicians as a group are better able to learn the structural rules of a new musical system than non-musicians. However, to lump musicians into a single group masks the large variation in pedagogy and performance demands. The current study was interested in the impact different types of musical training may have on the ability to learn from the auditory environment. Specifically, the current study considered musicians who have received either classical or jazz training. A total of 20 jazz students and 20 classical music students were recruited from Berklee College of Music and Boston Conservatory at Berklee. Participants completed a statistical learning task that was adapted from Mandikal Vasuki and colleagues (2016) and was based on a classical statistical learning task developed by Saffran (1999). Overall, both groups were above chance for their test performance indicating that they were able to learn the structural regularities embedded in the auditory stream. When considering musical training, it was found that jazz musicians performed significantly better than classical musicians at test. These results suggest that differences in practice strategies, listening skills, and performance demands between the genres of classical and jazz music may impact auditory learning and provide further evidence that types of musical training may have differential impacts on the system's ability to process and learn from the auditory environment.

Author e-mail: eknowles@berklee.edu

[Back to Poster Session](#)

15 **The Stability of the Speech-to-Song Illusion: Individual Differences**

Rodica R. Constantine*

University of Nevada, Las Vegas

Erin E. Hannon

University of Nevada, Las Vegas

Music and language are easily distinguishable for the average listener despite sharing many structural and acoustic similarities. In the Speech-to-Song (STS) illusion, multiple repetitions of a natural spoken utterance can give rise to a perceptual switch wherein the stimulus begins to sound like song to the listener. Although prior work suggests that both musicians and non-musicians experience the STS illusion (Vanden Bosch der Nederlanden et al. 2015a), some evidence shows that relative to non-musicians, musicians may require more repetitions to experience the STS illusion (Besson, Chobert & Marie, 2011; Dalla Bella, Peretz, & Aronoff, 2003; Falk et al., 2014), perhaps due to more robust musical representations or more refined selectivity of an auditory signal as speech- or song-like. In our study, we measured the STS illusion by presenting listeners with stimuli known to elicit the STS illusion and asking them to rate the degree to which each repetition sounded song-like. We also administered the Goldsmiths Musical Sophistication Index (Gold-MSI), a speech prosody test (PEPS-C), and a tonality test (from Corrigan & Trainor, 2015) to measure individual differences in musical and linguistic processing. Although data collection is ongoing, initial results suggest that musicality (reflected in the Active Engagement subscale of the Gold MSI) negatively predicts the rate of STS transformation, as found in prior work. Individuals who gave higher ratings overall also tended to get higher scores on the musicality and music training subscales of the Gold-MSI as well as higher scores on the speech prosody test. These findings provide initial support for the possibility that individual differences in musical skill, music aptitude, and sensitivity to speech prosody may predict the experience of the STS illusion.

Author e-mail: constr3@unlv.nevada.edu

[Back to Poster Session](#)

16 Misophonic Experience and Musicality

Sivan Barashy	UNLV
Solena Mednicoff	UNLV
Rodica Constantine*	UNLV
Stephen Benning	UNLV
Joel Snyder	UNLV
Erin Hannon	UNLV

Many musicians report experiencing misophonia, a condition that entails strong aversive emotional reactions to specific sounds (e.g., chewing, breathing, tapping) and avoidance of environments containing these triggers. An open question is whether individuals who are more likely to experience misophonia are also more likely to experience pleasurable auditory affective responses to stimuli such as music and autonomous sensory meridian response (ASMR). To address this question, we administered a survey to 200 participants, measuring misophonia sensitivity using the Amsterdam Misophonia Scale (A-MISO-S) and musicality using The Goldsmiths Musical Sophistication Index (Gold-MSI). We also measured momentary reactions to videos that were selected to induce experiences of misophonia, ASMR, and musical chills, by asking subjects to press a button every time they had an emotional reaction (whether positive or negative) and obtaining valence and arousal ratings for each video.

Participants' A-MISO-S scores were positively correlated and significantly predicted emotional reactions to negative (misophonic) videos, as well as emotional reactions to affectively positive (ASMR) videos. Emotional reactions to frisson videos and musicality scores (Gold-MSI overall score) predict both negative and positive emotional reactions, although musicality is not as strong of a predictor. This finding suggests that musical training and musicality might contribute to the enhanced responsiveness to unpleasant and pleasant sounds more generally.

Author e-mail: constr3@unlv.nevada.edu

[Back to Poster Session](#)

17 **Are Melodies a Useful Mnemonic Cue for Word List Recall?**

Teodor Zelenskyy

University of San Diego

Connor Warman

University of San Diego

Laura M. Getz*

University of San Diego

There is a colloquial belief that melodies can serve as an effective mnemonic cue, helping to remember information that would otherwise be harder to recall. However, research is mixed on the effectiveness of using music to enhance learning and remembering, based on the nature of the information to be remembered and the familiarity of the melody used. In the current study, we tested the effects of unfamiliar and familiar melodies on the memorization of unrelated word lists compared to a spoken control condition. Participants heard eight word lists with 15 two-syllable words either spoken, sung to a familiar melody, or sung to an unfamiliar melody. After hearing each list, they were asked to freely recall the words in any order and were then asked to perform an old-new recognition task. We found that participants performed best in the spoken word condition and worst in the familiar melody condition, suggesting that adding melodies to the word lists diverted attentional resources away from the memorization task, regardless of melody familiarity. However, when asked to report the name of the tune for the familiar melodies, only 14% were correctly recognized. Therefore, we are currently in the process of re-running the experiment (in the lab as opposed to online using mTurk) with different melodies that have been piloted to be more familiar to participants. Doing so will allow us to determine whether familiar melodies can in fact provide a mnemonic enhancement effect compared to unfamiliar melodies, or whether all melodies distract from the word recall task. These results will thus help to identify potential circumstances where the colloquial belief of music aiding memory holds true.

Author e-mail: lgetz@sandiego.edu

[Back to Poster Session](#)

18 **Perceptual evaluation of a novel method for synthesizing musical sounds**

Michael D. Hall*

James Madison University

Traditional tone synthesis methods typically obtain natural results without regard for computational load or implementation, and simplification generally produces artificial results. A promising alternative, Spectral Dynamic Synthesis (SDS; Beauchamp & Horner, 1995; Beauchamp, 2018), involves wavetable synthesis that reduces parameters by indexing a set of spectral envelopes according to perceptually relevant values, their spectral centroids. The current investigation assessed whether SDS tones are perceptually comparable to natural recordings for two instruments, trumpet (Experiment 1) and saxophone (Experiment 2).

Listeners completed tasks involving original and resynthesized versions of tones at F3, A#3, D4, F4, A#4, D5, and F5 pitches. 2AFC discrimination trials paired original and resynthesized versions of a tone, and listeners indicated which was resynthesized. 4IAX discrimination evaluated sensitivity to acoustic differences between versions, with listeners indicating which tone pair contained the resynthesized tone. Finally, listeners rated each tone's relative goodness for the instrument.

For trumpet, goodness ratings for resynthesized tones rivaled or exceeded the originals. In the 2IFC task mean sensitivity (d') was uniformly poor across pitches (> 1 only at A#4), meaning listeners couldn't distinguish which tones were artificial. Acoustic differences between versions were readily perceived ($d' > 1$) in 4IAX discrimination, though performance reduced slightly at higher pitches. Thus, resynthesis quality, while acoustically different, was excellent. In contrast, for saxophone goodness was typically higher for original tones. Listeners reliably distinguished ($d' > 1$) resynthesized tones in the 2AFC task (performance reduced slightly at higher pitches), and d' again remained high across pitches in the 4IAX task.

Acoustic analyses revealed more variable spectral envelopes over time for saxophone, with SDS producing harmonic amplitudes that deviated more from original values. Thus, SDS might most effectively approximate predictable (e.g., low- or band-pass filtered) spectral envelopes shapes. This implies that spectral centroids might not consistently predict detailed envelopes containing formant-like resonances.

Author e-mail: hallmd@jmu.edu

[Back to Poster Session](#)

19 **Distance perception of objects using visual-to-auditory sensory substitution: comparison of conversion methods based on sound intensity and envelope modulation**

Camille Bordeau*	Université Bourgogne Franche-Comté – Dijon, France
Florian Scalvini	Université Bourgogne Franche-Comté
Cyrille Migniot	Université Bourgogne Franche-Comté
Julien Dubois	Université Bourgogne Franche-Comté
Maxime Ambard	Université Bourgogne Franche-Comté

Visual-to-auditory sensory substitution devices are assistive devices for the blind that convert visual images into soundscapes by converting visual features into auditory cues. To convey spatial information with sounds, several sensory substitution devices use a Virtual Acoustic Space using Head Related Transfer Functions recordings to synthesize natural acoustic cues used for sound localization. However, these recordings are often acquired from one or a few sound source distances. Consequently, additional acoustic cues are often added to the signal to convey information about distance

In this study, we investigated the early-stage ability to judge the relative distance of two objects using soundscapes provided by a visual-to-auditory sensory substitution device. Our experiment aimed to compare methods to convey distance using audio signal intensity and envelope modulation of the sound.

Blindfolded participants had to discriminate the distance of two virtual targets using soundscapes. We tested two visual-to-auditory conversion methods for distance: one using sound intensity modulation, and one using the same sound intensity modulation but associated with an audio signal envelope modulation.

Results show that participants perceived relative distance of objects using soundscapes provided by both conversion methods with a better performance while coupling the intensity modulation with an envelope modulation. These results provide evidence that modulation of audio signal envelope is an interesting method to emphasize information relative to distance in visual-to-auditory sensory substitution devices.

Author e-mail: camille_bordeau01@etu.u-bourgogne.fr

[Back to Poster Session](#)

20 **Cochlear implant users and normal-hearing listeners have contrasting hemodynamic changes during auditory beat processing despite similar behavioral response**

Samantha R. O'Connell*	Keck School of Medicine, University of Southern California (USC), Los Angeles
Ali Rahimpour Jounghani	University of California, Merced
Julianne M. Papadopoulos	Thornton School of Music, USC, Los Angeles
Heather Bortfeld	University of California, Merced
Raymond L. Goldsworthy	Keck School of Medicine, USC, Los Angeles

Introduction: Cochlear-implant (CI) users have poor pitch perception which makes music listening and speech comprehension difficult; however, they have similar rhythm perception to normal-hearing (NH) individuals. Currently, it is unclear 1) what brain regions are involved in processing rhythmic timing information in CI users and 2) if beat-related cortical activation patterns in CI users are like NH listeners.

Methods: Using a robust regression model, hemodynamic changes were measured using function near-infrared spectroscopy (fNIRS) in motor and auditory cortices of right-handed CI users and NH listeners during a beat synchronization finger tapping task. Pseudo-replicating a previous experiment by Rahimpour et al. (2020), participants were asked to either tap along with (synchronization) or in between (syncopation) 1Hz isochronous auditory tones. After 15 seconds, participants were asked to maintain the beat in both conditions after auditory tones ceased (continuation).

Results: CI users show greater activation of supplementary motor area (SMA) and premotor cortex (pMC) in synchronized pacing and continuation conditions compared to syncopated conditions. In syncopated pacing and continuation conditions, hemodynamic activity migrated from motor planning regions (SMA/pMC) to frontal areas. Conversely, NH listeners exhibit heightened hemodynamic activation in frontal regions during synchronized continuation and pacing conditions changing to motor planning regions during syncopated tapping. Interestingly, behavioral responses do not reflect these differences: reaction times between the CI and NH listeners are similar.

Conclusion: Our results indicate that beat perception involves contrasting brain regions in CI and NH listeners. CI listeners seem to rely on motor planning regions to synchronize to a beat and recruit attention-related regions when tapping between beats - a naturally more difficult task. Differently, NH listeners may engage frontal, attention-related regions more when trying to tap accurately along with a beat but may use motor planning regions - known to be linked to auditory beat processing - when tapping in syncopation.

Author e-mail: sroconne@usc.edu

[Back to Poster Session](#)

21 **Restrictive eating tendencies and variability in audiovisual integration**

Annika Beatteay*

University of New Brunswick

Jonathan Wilbiks

University of New Brunswick

Although there is a significant body of research surrounding the cognitive deficits that often correspond with restrictive-type eating disorders, little research has been conducted examining audiovisual integration abilities in this population. Given the abnormalities noted in other forms of multisensory integration in this population, we hypothesized that individuals with more restrictive tendencies and more severe overall ratings of illness would have increased difficulty appropriately perceiving synchrony of sights and sounds, in addition to a bias for local processing. Participants were required to complete two self-report questionnaires that assessed the presence and severity of symptomatology (EDE-Q & EAT-26). Participants also completed the Navon task to assess whether individuals were more likely to focus on local versus global aspect of the stimulus as seen in previous eating disorder research. Finally, participants completed a version of the Sound Induced Flash Illusion testing both fission and fusion effects to assess the accuracy of audiovisual integration and the width of the temporal binding window. There does not appear to be a significant correlation between local precedence bias and more severe symptoms. While investigating the width of the temporal binding window, an independent samples t-test was conducted with the clinically significant EDE-Q and the SIFI illusion. In individuals with clinically relevant EDE-Q scores, a small but significant difference was found in Fission trials. This indicates that individuals with more significant eating disorder behaviours have a wider temporal binding window. Additionally, a correlational analysis was conducted between the Restraint subscale of the EDE-Q and the SIFI task. This analysis indicated a weak relationship between performance on Fission trials and the severity of restrictive behaviours endorsed on the EDE-Q. This abnormally wide temporal binding window in individuals with severe restrictive-type symptoms paves the way for new, potentially very effective methods of treatment through recalibration of the binding window.

Author e-mail: abeattea@unb.ca

[Back to Poster Session](#)

22 **Behavioral Markers of Age-Related Hearing Loss Revealed through the
Attentional Network Task**

Gennadiy Gurariy*	Medical College of Wisconsin; Marquette University
Sarah Mleziva	Medical College of Wisconsin
Lauren Elliot	Medical College of Wisconsin; Marquette University
Samiah Ziadeh	Medical College of Wisconsin; Marquette University
Kristin Kozlowski	Medical College of Wisconsin
Michael S. Harris	Medical College of Wisconsin
Adam S. Greenberg	Medical College of Wisconsin; Marquette University

Age-related hearing loss is an escalating public health concern given the growth of senior populations. The relationship between hearing loss and cognitive decline is well-recognized based on epidemiological data. Two hypotheses currently predominate regarding the nature of this relationship. The sensory deprivation hypothesis states that potentially irreversible brain changes arise from decreased sensory stimulation. Alternatively, the information degradation hypothesis states that the ambiguity of the auditory signal necessitates recruitment of additional resources that are diverted from other cognitive processes. Unlike the sensory deprivation hypothesis, the information degradation hypothesis suggests that cognitive decline should be reversed following restoration of audition. To better understand the link between hearing loss and cognitive decline, we recruited patients with untreated hearing loss who qualify for but have not yet undergone cochlear implant (CI) surgery (PreCI) as well as normal hearing controls. To assess cognitive function, we utilized two behavioral paradigms: (1) the N-back task, which measures working memory, and (2) the Attentional Network Task (ANT), which quantifies the alerting, orienting and distractor filtering dimensions of attentional control. Our results did not reveal any differences in working memory performance between the two groups. However, the ANT results do suggest differences between the two populations with the control group showing a stronger alerting effect and the PreCI group showing a stronger orienting effect. Interestingly, no differences were observed for distractor filtering. These preliminary results suggest that PreCI patients show a different pattern of behavior relative to healthy controls on some measures of cognitive function. Future work will involve (1) collecting additional data from PreCI participants, (2) collecting data from newly implanted CI users, (3) measuring behavioral performance longitudinally of the PreCI cohort after CI surgery, and (4) using fMRI to examine changes in brain function associated with hearing loss driven cognitive decline.

Author e-mail: ggurariy@mcw.edu

[Back to Poster Session](#)

23 **Acoustic profile of French emotional sentences uttered with French and Quebecois accents**

Amélie Rolinat*

BRAMS, CRBLM, UdeM

Simon Rigoulot

BRAMS, CRBLM, UdeM, UQTR, CogNAC

Speech prosody, i.e. variations of the tone of voice when speaking, plays a key role in social interactions by bringing important information related to identity, emotional state and geographic origin. For example, foreign accents have a strong impact on how speech is recognized, with important consequences on social evaluations. However, it is less clear if this impact persists in the context of regional accents that constitute more subtle variations of the vocal signal. Consequently, the objective of this study was to understand how french-speaking individuals from different regions (France, Québec) recognize emotional sentences uttered by people originating from the same region or not. In a first step, we built a bank of emotional sentences uttered with regional French and Quebec accents in order to study the acoustic profile of Quebecers and French when conveying emotions. Forty sentences depicting an emotional situation (anger, happiness, surprise, sadness, jealousy, shame and pride) such as 'What a beautiful day' or 'You are still late', were interpreted by 6 Quebec and 6 French actors (3 women for each group). We analyzed some acoustic parameters of the sentences (F0, intensity, energy profile and speech rate) related to speech prosody and found significant differences as a function of the origin of the actors and the conveyed emotion. Among others, we found that mean F0 was higher when expressing anger in the French than in the Quebecers and that French actors were overall speaking with a higher intensity than the Quebecers when expressing shame. Following this, we aim to validate the use of these sentences through an online study where participants will have to evaluate sentences on the quality of accents and emotions. The database will further serve for a neuroscientific experiment where we will investigate the brain responses to sentences uttered with same and different regional accent.

Author e-mail: amelie.rolinat@umontreal.ca

[Back to Poster Session](#)

24 **Voice-induced synesthesia, emotion and plasticity**

Cathy Lebeau* UQÀM
Gabrielle Lebeau UQÀM
François Richer UQÀM

Voice-induced synesthesia are atypical perceptions triggered by specific voices in relation to their identity or features. We report four cases suggesting early experiences and emotions play important roles in voice-induced synesthesia. Case 1 reported that all voices evoked visual patterns linked to the timbre and prosody of voices but when the voices evoked emotional reactions, they triggered an additional colored halo. Case 2 reported that some female voices triggered texture sensations as well as an impression of emotional blending and heart churning. Case 3 experienced transient blindness at an early age after which she recovered some vision for low frequencies. Several voices evoked sensations of textures. While beautiful and touching voices made textures more vivid, nasal voices evoked aversive sensations. Case 4 reported that numerous voices of celebrities evoked tastes and some tastes could be linked to specific childhood experiences. These cases suggest that emotions can modulate voice-induced synesthesia. They also suggest that early plasticity and childhood experiences can induce synesthetic associations.

Author e-mail: lebeau.cathy@courrier.uqam.ca

[Back to Poster Session](#)

25

It's not who you are, or what you say, but how many people say it: A multi-talker study of vocal confidence

Katie Asaro*

Kent State University

Bradley J. Morris

Kent State University

Jennifer M. Roche

Kent State University

Speaking with confidence is an integral aspect of effective communication both at a personal and professional level. Socio-indexical cues to gender and vocal affect often interact and sometimes lead listeners to make differential judgements of affective intent based on the gender of the speaker. Previous research suggests that rising intonation is a common cue that both men and women produce to communicate lack of confidence, but listeners are more accurate in identifying speaker intent when the cue is produced by a woman. It is less clear if this cue is truly gender specific and modifies a listener's interpretation of affective intent. We conducted a study in which we evaluated listener perception of speaker intent to communicate confidence (or lack thereof) by many speakers (4 women; 4 men) to determine if the salience of an intonational cue to confidence comes from its interpretation or is modified by socio-indexical cues to gender. We hypothesized that if confidence is interpreted based on the interaction between the socio-indexical cues of gender and confidence, then regardless of number of instances and types of confidence cues (e.g. rising intonation) produced across gender, the men speakers should always be perceived as more confident than women when rising intonation is produced. Results confirmed that when listeners listened to an individual woman and man talker, they were more likely to differentiate their judgements based on gender. However, when listeners judged multiple talkers, listeners were less likely to differentiate their interpretation of confidence based on the interaction between gender and affective intonation. This suggests that when multiple talkers produced the same type of cue, listeners deemed the relevant cue as intonational and not an interaction between gender and affect. We may, therefore conclude, that it is not who you are or how you say it, rather it is how many people say it like you do.

Author e-mail: kasaro@kent.edu

[Back to Poster Session](#)

26 **Sensorimotor Processing of Vocal Emotion**

Emma Greenspon*	Monmouth University
Adriana Zefutie	Monmouth University
Marissa Lagera	Monmouth University

In the current study, participants were presented prerecorded pseudosentences spoken by male and female speakers that portrayed one of six emotions: anger, happiness, fear, sadness, disgust, and neutral emotion. Participants then mentally rehearsed the sentence while completing one of four interference conditions: no interference, silent articulation (articulatory interference), quiet humming (phonatory interference), and finger tapping (non-vocal motor interference). In addition, participants completed a singing task in which they imitated four-note novel melodies and a simple pitch discrimination task which involved judging whether a comparison pitch was lower or higher than a target pitch. The study was completed online using the platform Finding Five. We ran a 4 (interference condition) x 6 (emotion type) repeated measures ANOVA on unbiased accuracy rates for a sample of 42 participants. A significant effect of emotion type ($p < .001$) revealed that, as expected, sentences portraying disgust were recognized with the lowest accuracy ($M = .33$) and sentences portraying fear were recognized with the highest accuracy ($M = .79$). There was a significant interaction between emotion type and interference condition ($p < .01$), but no main effect of interference condition ($p = .84$). The interaction was driven by a single significant contrast: articulatory interference ($M = .46$) during rehearsal disrupted vocal emotion recognition relative to finger tapping ($M = .62$) for sentences portraying sadness. We averaged unbiased accuracy rates across emotion types and interference conditions to determine whether singing accuracy or pitch discrimination accuracy was a stronger predictor of overall vocal emotion recognition. Hierarchical linear regression replicated our previous finding that singing accuracy was the only unique predictor of vocal emotion recognition while controlling for pitch discrimination accuracy. We interpret the results of the hierarchical linear regression as evidence that sensorimotor processing of the vocal system relates to vocal emotion recognition.

Author e-mail: egreensp@monmouth.edu

[Back to Poster Session](#)

27

**An acoustic analysis of probabilistic language used during COVID-19
White House Press Briefings**

Yin Zhang*

Kent State University

Bradley J. Morris

Kent State University

Jennifer M. Roche

Kent State University

COVID-19 has impacted every corner of the world for two years, yet, the public has not unanimously recognized the severity and impact of the pandemic. Why do people have different opinions toward the pandemic? One reason may be how health-related information has been communicated. Research on pragmatics and communication style suggests that probabilistic language (e.g., hedging) may impact the interpretation of facts and myths. Prior research suggests that hedging may make a statement more believable — because it communicates humbleness. However, scientists and politicians may hedge for very different reasons. For example, scientists may hedge to communicate the probabilistic nature of science, while politicians might hedge to persuade. In the current study, we evaluated perceptual differences in the social judgments of the hedging word ‘might’ by public figures. In this naturalistic production/perception study, we selected and digitally extracted fifty-three audio files where public figures (i.e., the Press, Political Advisors, and Public Health Officials) produced the hedge word ‘might’ during COVID-19 White House Press Briefings over the last two years. These audio files were then presented to listeners, who made social judgments about the confidence (or lack thereof) of the speaker’s might utterances. Results showed that public figures spoke with variable acoustic characteristics that listeners used to perceive confidence. The acoustic data suggested that the press, political, and public health officials varied their acoustic cues when hedging. Preliminary data suggest that listeners were far more likely to use the acoustic cues to predict confidence when produced by a public health official. More specifically, when public health officials produced rising intonation ($p < .001$) and shorter speaking durations ($p = .02$), listeners used these cues as indicators to make confidence judgments. Findings from this study provide the groundwork to investigate the different contexts in which people hedge in public communication.

Author e-mail: yzhan108@kent.edu

[Back to Poster Session](#)

28 **Assessing the Effects of “Native Speaker” Status on Classic Findings in
Speech Perception**

Katrina Sewell

Yuxin Lin

Violet A. Brown

Julia F. Strand*

Carleton College

Carleton College

Washington University in St. Louis

Carleton College

One of the goals of psychological research is gaining a better understanding of human behavior, yet many studies make use of samples that do not represent the human population generally (e.g., undergraduate students). Limited sampling is common across subdisciplines of psychology, but this issue is exacerbated in research on spoken language, by constraining the sample to “native English Speakers.” Not only is the “native English speaker” criterion poorly defined, but it also excludes historically underrepresented groups from speech perception research, often without attention to whether this exclusion is likely to affect study outcomes. The purpose of this study is to empirically test whether and how using different inclusion criteria (“native English speakers” vs. “non-native English speakers”) affects several well-known phenomena in speech perception research. Five hundred participants completed word (N = 200) and sentence identification (N = 300) tasks in quiet and in moderate levels of background. Here we present data indicating that multiple classic findings in speech perception research—including the effects of noise level, semantic context, and lexical density on speech intelligibility—persist regardless of whether participants meet the commonly-used criteria used to define “native English” speaking status. However, the magnitude of some of these effects (and the interactions among them) differed across participant groups.

Author e-mail: jstrand@carleton.edu

[Back to Poster Session](#)