APCAM 2024

23rd Annual Auditory Perception, Cognition, & Action Meeting

Thursday, November 21st, 2024

New York Marriott Marquis

New York, NY

8:00am - 5:00pm

23rd Annual Auditory Perception, Cognition, and Action Meeting

Welcome to the 23rd annual Auditory Perception, Cognition, and Action Meeting (APCAM 2024)! 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."

APCAM is a unique meeting that blends basic and applied research from different theoretical perspectives and numerous types of auditory stimuli (including speech, music, and environmental sounds). The continued flourishing of APCAM is 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 2024 are eligible to submit a brief report for consideration in a special issue of *AP&C*. Further information on this opportunity is available from the editors of *AP&C*, Michael Hall (hallmd@jmu.edu) and Mike Russell (mirussell@bellevue.edu). In addition, we encourage you to submit your other work on auditory science to *AP&C*.

APCAM is 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 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 65th Annual Meeting of the Psychonomic Society, APCAM is indebted to the Psychonomic Society for material support. This year we welcome Gorilla experiment building software as a new sponsor, and thank all of our new and returning sponsors for making our conference possible.

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 by **David Poeppel**; 19 spoken sessions; and 32 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 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 2024 Organizing Committee

Timothy L. Hubbard (Chair)

J. Devin McAuley

Kathleen C. McCulloch

Kristopher J. Patten

Peter Q. Pfordresher

Hannah Shatzer

Sponsors and Donors

Thanks to our sponsors and donors!

Gold Sponsors

PSYCHONOMIC SOCIETY®

https://www.psychonomic.org/default.aspx

Psychology
Carnegie Mellon

https://www.cmu.edu/dietrich/psychology/

Silver Sponsors

FINDING FIVE

https://www.findingfive.com/

www.gorilla.sc

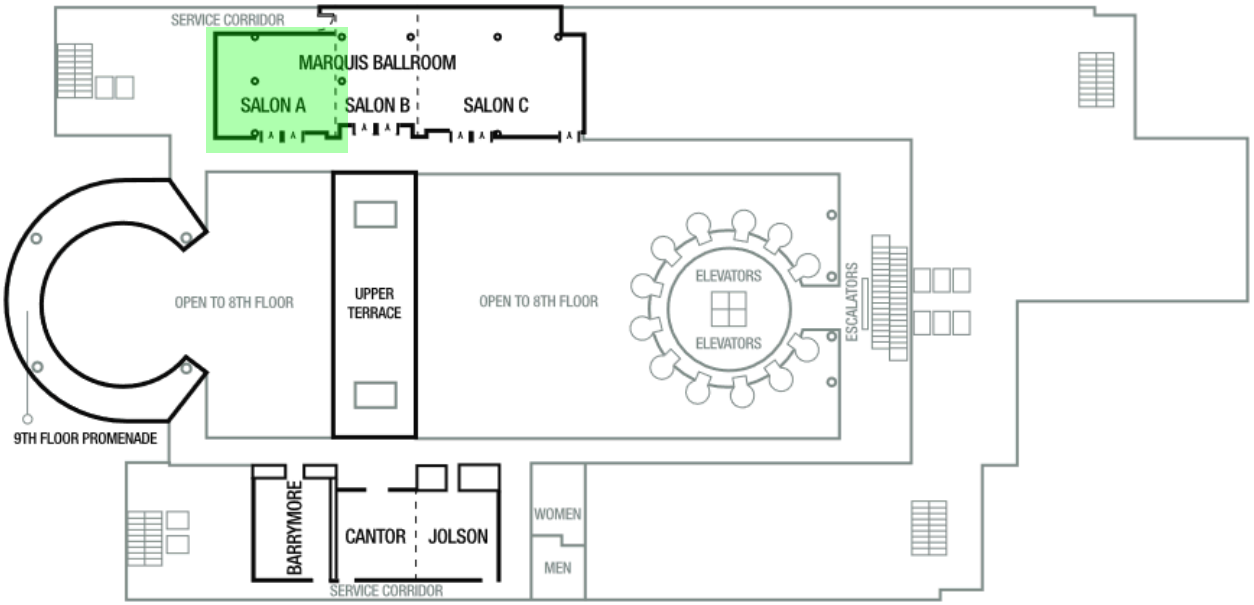
https://gorilla.sc/

Member donations were made by:

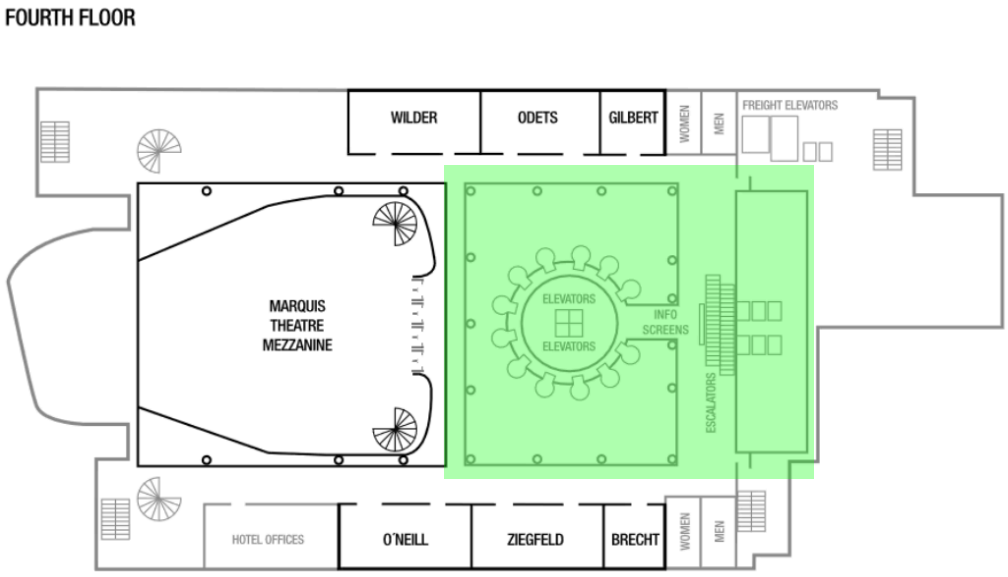
- Melissa Baese-Berk
- Annabel J. Cohen
- Leah Fostick
- Laura Getz
- Laurie Heller
- Erica Knowles
- Psyche Loui
- Michael K. McBeath
- Jenny Roche
- Chad Rogers

Map

Spoken Sessions - Marquis Ballroom, Salon A
Ninth Floor



Poster Session - Foyer
Fourth Floor



Full Schedule

8:00	**Registration**	
8:10	**Welcoming Remarks**	
8:15	**Music**	
8:15	The sound of manufactured music: Reviewing the role of artificial stimuli in music cognition research	**Douglas A. Kowalewski**, U. at Albany, SUNY
8:30	Geometric relationships across notes constitute a primitive in melodic processing	**Omri Raccah**, New York University
8:45	Effects of select musical modes on perceived emotion of written stimuli	**Laurel Aronian**, St. Luke's School
9:00	Dissociation of musical pitch processing in perception and production: A case study	**Peter Q. Pfordresher**, U. at Buffalo SUNY
9:15	Timbre descriptors in speech and song expression	**Lena Heng**, U. of Prince Edward Island
9:30	**Break**	
9:45	**Cross-Modal & Multi-Modal Perception**	
9:45	Distinct impairments in sensorimotor integration mediate auditory hallucinations in schizophrenia	**Xing Tian,** NYU Shanghai
10:00	Seeing sounds, feeling shapes: A tri-modal analysis of cross modal correspondence and synaesthesia	**Luke Lucas**, London Metropolitan U.
10:15	Naturalistic multimodal spatial interactions	**Laurie M. Heller**, Carnegie Mellon U.
10:30	Improving distance perception in virtual reality using spatially incongruent real-world sounds	**Maggie McCracken**, University of Utah
10:45	Effects of meter-rhythm complexity and stimulus sound intensity on the vigor of synchronous walking	**Gregory D. Shay**, Independent Researcher

11:00	**Break & Poster Set-Up**	
11:15	**Poster Session**	
12:30	**Lunch**	
1:30	**Keynote**	
	Rhythms and Algorithms: From Vibrations in the Ear to Abstractions in the Brain	**David Poeppel**
2:00	**Auditory Scenes & Timing**	
2:00	Did I do that? Investigating objective auditory-motor agency in younger and older adults	**Alexis Basciano**, Arizona State University
2:15	Auditory perception of the material composition of everyday objects	**Aiden Iveris**, UC San Diego
2:30	Positive feedback, negative impact: The role of feedback in auditory temporal order judgment	**Leah Fostick**, Ariel University
2:45	An empirical study of synchronization behaviour and experience in human-machine trios	**Bavo Van Kerrebroeck**, McGill University
3:00	**Break**	
3:15	**Speech & Language**	
3:15	The Morgan Freeman effect: The advantages and disadvantages of voice familiarity when listening to speech while distracted	**Manda Fischer**, U. of Western Ontario
3:30	Intrinsic voice memorability affects memory for spoken content	**Cambria Revsine**, University of Chicago
3:45	Contributions of distributed training and semi-supervised learning to improvement on a non-native phonetic classification task	**Beverly A. Wright**, Northwestern University
4:00	Decoding semantic categories for spoken words and images from EEG data	**Emma Karn**, Fordham University
4:15	On the role of orthography and images in early foreign language learning: An ERP study	**Mathew Cieśla**, Northumbria University
4:30	**Meeting of the Membership**	

Poster Session

1-10	Speech & Language	
1	Engaging Stories for the Study of Attention and Audition: An introduction to the ESSAA database	**Lauren Petley**, Clarkson University
2	Intelligibility of medically related sentences in quiet, speech-shaped, and hospital noise	**Sarah White**, University of Chicago
3	Correlation between speech movement planning, task complexity, and beta frequency activity in normal adults	**Mohsen Parsa**, Arizona State University
4	Equivalent processing of self-relevant information in autism: Evidence from the cocktail party phenomenon	**Jan Philipp Röer**, Witten/Herdecke U.
5	Navigating uncertainty: The influence of cognitive and non-cognitive factors on the interpretation of provisional language in science communication	**Shruti Kate**, Kent State University
6	A meta-analytic perspective on characteristics of brief speech alerts	**Michael A. Nees**, Lafayette College
7	Assessing AI generated speech: The impact of gender, (in)directness, and intonation on the perception of rudeness	**Jacob Miller**, Kent State University
8	Physical exertion as a novel, in-the-moment index of listening effort	**Carson Rumble-Tricker**, University of Guelph
9	English-Spanish bilinguals' perception of ambiguous speech sounds changes based on task language	**Gabriella Thomas**, University of San Diego
10	Difficulty of speech-in-speech recognition: How does listening environment modulate informational masking?	**Madelene Tavarez**, Penn State University
11-18	**Music**	
11	From performer to listener: Understanding emotional perception in music	**Emma Ning**, U. of Illinois Chicago

12	Statistical learning: How contingent regularities influence learning, memory and preference for melodic pairs over the time-course of learning	**Pierce C. Johnson**, U. at Albany, SUNY
13	Exploring the impact of auditory manipulations on piano learning: Sensory deprivation, imagery, and skill transfer	**Renan Moreira Madeira**, U. of Illinois Chicago, UFRGS
14	The Musician Recognition Task: A new measure assessing musical exposure as it relates to musicality	**Veronica Kando**, University of Maryland
15	Melodies are not a useful mnemonic cue for immediate word recall regardless of familiarity	**Isabella Lea Ramirez**, University of San Diego
16	Is adolescence a sensitive period for acquisition of musical knowledge? Behavioral evidence in the light of pop music style change analysis	**Annabel J. Cohen**, U. of Prince Edward Isl.
17	Developing written music literacy: The relationship between working memory and melodic dictation ability	**Erica R. Knowles**, Berklee College of Music
18	Absolute Pitch among non-global possessors - Timbre, categorization and musical pleasure	**Ébano Resende de Souza**, Austrian Acd. of Sciences

19-24	**Auditory Scene Analysis**	
19	Impact of sound source motion on the perception of occlusion and disocclusion	**Mike Russell**, Bellevue University
20	Context-dependent auditory inference in environmental sound recognition	**Keland Moore**, University of Iowa
21	Pupil-linked arousal tracks adaptive auditory belief updating in spatially and temporally dynamic environments	**Roman Fleischmann**, Austrian Acd. of Sciences
22	Investigating cognitive disruptions in misophonia: The impact of trigger sounds on working memory performance	**Michael A. Tollefsrud**, Kansas State University
23	Non-auditory impacts of wearing hearing protection: concurrent sensorimotor tracking and choice to wear	**Matthew G. Wisniewski**, Kansas State University
24	Test of a selective entrainment hypothesis using concurrent tone sequences	**Toni Smith**, Michigan State U.

25-27	Methodological Issues	
25	Perceptual discrimination of changes in bit-depth resolution	**Michael D. Hall**, James Madison U.
26	Perceptually-based considerations regarding the digital representation of amplitudes	**Michael D. Hall**, James Madison U.
27	Strategies for fitting response time data in psychoacoustic experiments: Insights from a noisy exemplar approach	**Nathan F. Gillespie**, U. at Albany, SUNY
28-29	**Cross-Cultural Perception**	
28	Culture influences multisensory emotion perception in bicultural bilinguals	**Viorica Marian**, George Fox University
29	The role of basal ganglia in Chinese garden-path sentence comprehension processing	**Yingying Tan**, Northwestern University
30-32	**Cross-Modal Perception & Synesthesia**	
30	Shocking results: Impacts of auditory-tactile training on auditory perceptual learning	**Chelsea Joyner**, Emporia State U., Kansas State U.
31	Emotions and altered consciousness in voice-induced synesthesia	**Cathy Lebeau**, U. du Québec à Montréal
32	Congruency advantage in multisensory processing is task-specific	**Kayla Phan**, University of San Diego

Diversity, Equity, and Inclusion Committee Updates

Members: Laura Getz (Chair), Laurie Heller, CJ Joyner, Chad Rogers, and Hannah Shatzer

Membership Demographic Survey

The DEI committee is assessing the needs of our members and addressing concerns related to diversity and equity within our community by completing a membership survey. The purpose of this survey is to gather information about the identities, backgrounds, and experiences of APCS members. This survey is anonymous, and data will be analyzed in aggregate to protect confidentiality of responders. These results will be shared with the DEI committee and APCS board to help assess the needs of the APCS community and establish actionable goals to support underrepresented scholars in auditory science. This survey will take approximately 10 minutes to complete. Answers to all questions are optional, but we would greatly appreciate your feedback!

Link here or scan the QR code: https://forms.gle/9rttRt23Vbe72xB7A

2024 Diversity Award Recipients

More information: https://apcsociety.org/dei-awards.html

Travel Award: $500 travel award for undergraduate or graduate students from underrepresented populations broadly defined.

- Aiden Iveris (University of California San Diego)

Membership Award: free registration for participants with financial need regardless of rank.

- Cathay Lebeau (Université du Québec à Montréal)
- Shruti Kate (Kent State University)
- Sarah White (University of Chicago)
- Mohsen Parsa (Arizona State University)
- Roman Fleischmann (Austrian Academy of Sciences)
- Veronica Kandro (University of Maryland)
- Gregory Shay (Retired Independent Scientist)
- Nathan F. Gillespie (University at Albany, SUNY)

Talk Abstracts

Music (8:15-9:30)

8:15 The sound of manufactured music: Reviewing the role of artificial stimuli in music cognition research

Douglas A. Kowalewski* University at Albany, SUNY

Having participants listen and react to musical stimuli is one of music cognition's foundational methods. Whereas most researchers have used stimuli adapted from existing musical traditions in such work, others have incorporated artificial stimuli (i.e., stimuli generated specifically for research that are not borrowed from any existing sociocultural musical system) into their research designs. No review of this growing literature on artificial stimuli exists, leaving open the question of how useful they are in helping formulate and test research questions in music cognition. To this end, a systematic narrative review of empirical studies utilizing artificial musical stimuli (N = 52) was conducted. Comparing these studies to analogous works involving conventional stimuli in the areas of music preference, music learning, and musical emotion revealed the power of artificial stimuli in reducing concerns regarding stimulus familiarity and musical enculturation confounds, as well as in refining theory by reiterating the need for the development of composite predictive models in music cognition. Emerging evidence also indicates that artificial stimuli can help address similar questions relating to musical memory, tonal affinity, and imagination. This review additionally suggests that using artificial stimuli has the potential to help clarify the neurological and sociocognitive factors surrounding music listening behavior more generally. Furthermore, because artificial stimuli are not tied to any existing sociocultural musical system, they may also lead to more generalizable findings. This enhanced generalizability may also coincide with expanded inclusivity, toward which the field of music cognition (and psychology more generally) are increasingly striving. Other potential uses, benefits, and limitations of artificial musical stimuli are discussed, as well as recommendations for researchers seeking to utilize such stimuli in future studies. In sum, this review suggests how the conscientious utilization of artificial music stimuli can help advance music cognition into a more inclusive and theoretically sophisticated future.

Author e-mail: dkowalewski@albany.edu

Back to Schedule

8:30 Geometric relationships across notes constitute a primitive in melodic processing

Omri Raccah*	Yale University
Michael Seltenreich	New York University
Claire Pelofi	New York University
Fred Lerdahl	Columbia University
David Poeppel	New York University

How do the notes included in a melody affect its processing? The interactions and relationships among collections of notes are surprisingly intricate, and their impact on perception is not well understood. Theoretical accounts have suggested that musical sets - the relational structures resulting from any collection of notes - constitute a critical aspect of music perception. However, empirical work which directly engages with this foundational construct remains limited. We introduce a simple behavioral protocol to test how musical sets, and their geometric structure, influence melodic sequence processing. Based on data collected from hundreds of participants across three experiments, we demonstrate that certain sets wholly alter our sensitivity to note deviations in melodies. We further show that geometric properties, including the uniformity of note distributions, capture these effects across a variety of musical structures. Altogether, our results firmly position musical sets as a primitive representation in melodic processing and uncover geometric measures that account for their role in music perception.

Author e-mail: omri.raccah@yale.edu

Back to Schedule

8:45 Effects of select musical modes on perceived emotion of written stimuli

Laurel Aronian* St. Luke's School, New Canaan, CT

This study analyzes how musical modes affect the perception of emotion in written stimuli among high school students. The study builds on Bostwick and Seror (2018), who investigated emotions perceived from musical modes. Musical modes are a type of scale, or series of notes in fixed intervals, derived from the major scale. The Ionian mode is perceived as the most positive mode, while the Locrian mode is perceived as more emotionally ambiguous. In this study, Ionian and Locrian mode "melodies" were created with random notes and a consistent drone note underneath to establish the modes. To test how modes affected the way participants perceived the emotion of written stimuli, three trials were created consisting of three poems, randomly presented, with the "melodies" playing behind each poem. Control participants read the same poems without melodies behind them. After each trial, all participants rated each poem's perceived valence and arousal using Betella and Verschure's Affective Slider. This study also used a cacophony primer (a series of random notes) between trials to discourage negative primacy effects. Preliminary analysis with a MANOVA using musical mode as a between-subjects variable, poem type and arousal type as within-subjects variables indicated that there was a significant effect of musical mode on the combined dependent variable, $F(2,29)=4.04$, $p=.028$, $n^2=.218$. The combined effects of pleasure and arousal were significantly different, $F(1,29)= 5.76$, $p=.023$, $n^2=.166$. The control group had lower levels of both arousal and pleasure in all three poems combined when evaluating the mean pleasure and arousal values. These findings suggest that music has a positive effect on the perceived emotion of written stimuli regardless of whether Locrian or Ionian is utilized.

Author e-mail: laurelaronian@gmail.com

Back to Schedule

9:00 Dissociation of musical pitch processing in perception and production: A case study

Peter Q. Pfordresher*	University at Buffalo, SUNY
Madeleine R. Kates	Independent Researcher
David J. Vollweiler	University of Nevada at Las Vegas
Chihiro Honda	University at Buffalo, SUNY

We present results from a case, referred to as "Patient", who experienced severe alterations to neural and physical functioning following routine vaccinations for college. Whereas most of these changes led to deficits, Patient seemed to experience spared or even improved functioning in the domain of music. We report laboratory tests of these anecdotal reports. Patient completed a battery of measures that included singing accuracy, imitation of spoken and sung pitch patterns, music perception, mental imagery (for musical pitch, phonology, and visual scenes), and short-term memory (for musical pitch and spoken digits). Results from most tests bore out Patient's anecdotal experiences: She performed abnormally well in measures of singing accuracy, voice quality, pitch perception, and self-reported auditory imagery vividness. Patient was spared (normal functioning) for measures of pitch span and accuracy of auditory imagery. By contrast, Patient underperformed on digit span and visual imagery. Poor performance was also noted on imitation of spoken and sung pitch, possibly reflecting the reduced tonal salience of these stimuli. Overall, these results converge with other evidence for neural separation of musical pitch processing from other auditory and cognitive processes. The fact that these deficits were found in a case who did not experience tissue damage suggests that adverse physical reactions may lead to neural rewiring in the absence of damage.

Author e-mail: pqp@buffalo.edu

Back to Schedule

9:15 Timbre descriptors in speech and song expression

Lena Heng*	University of Prince Edward Island
Annabel J. Cohen	University of Prince Edward Island

Features of timbre have been shown to differentiate emotions expressed in music and speech. This study assessed the timbre features characterizing the full set of audio recordings of expressed spoken and sung emotions in The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) (Livingstone & Russo, 2018). RAVDESS is a validated database containing recordings from 24 professional actors, singing and speaking a set of emotions expressed in two semantically neutral English sentences. The song stimuli represent six emotions: neutral, calm, happy, sad, angry, and fearful; disgust and surprised are added to these six to comprise the speech stimuli. No prior study has comprehensively assessed the acoustic features that differentiate these emotion categories in RAVDESS. We performed an acoustic analysis of all audio recordings with the revised Timbre Toolbox (Kazazis et al., 2021) implemented in the MATLAB environment, obtaining 58 spectral, temporal, and spectrotemporal descriptors for each of the sung and spoken audio clips. Random forest, a supervised machine learning algorithm was used to identify the acoustic descriptors that best classify the emotions. A partial least squares discriminant analysis was also performed to better understand how a smaller number of components might group these descriptors. Results revealed several descriptors that carried the greatest weight. Spectral variation (representing the amount of change of the spectrum over time) and effective duration were similarly important for both expressed speech and song. However, only for song, did tristimulus 1 carried a high weight, and conversely, for speech, pitch was heavily weighted. These results complement prior evidence and speculation of the similarity between acoustic parameters underlying expression by speech and music of the same emotion. However, this comprehensive assessment of 58 timbral descriptors points not only to certain shared qualities in vocal expression, but also to differences specific to the modality of speech or song.

Author e-mail: lheng@upei.ca

Back to Schedule

Cross-Modal & Multi-Modal Perception (9:45-11:00)

9:45 Distinct impairments in sensorimotor integration mediate auditory hallucinations in schizophrenia

Fuyin Yang — Shanghai Jiao Tong University
Xing Tian* — NYU Shanghai

Distinguishing reality from hallucinations requires efficient monitoring of agency. It has been hypothesized that a copy of motor signals, termed efference copy (EC) or corollary discharge (CD), suppresses sensory responses to yield the sense of agency; impairment of the inhibitory function leads to hallucinations. However, how can the sole absence of inhibition yield positive symptoms of hallucinations? We hypothesize that selective impairments in functionally distinct signals of CD and EC cause the positive symptoms of hallucinations. In an EEG speaking experiment in schizophrenic patients with (AVHs) and without (non-AVHs) auditory hallucinations, we found that preparing to speak without knowing the contents did not suppress auditory responses in both groups. Whereas, preparing to speak a syllable enhanced the auditory responses to the prepared syllable in non-AVHs, but AVHs showed enhancement for unprepared syllables. 'Broken' CD plus 'noisy' EC causes erroneous monitoring on the imprecise internal auditory representation and yields hallucinations.

Author e-mail: xing.tian@nyu.edu

Back to Schedule

10:00 Seeing sounds, feeling shapes; A tri-modal analysis of cross-modal correspondence and synaesthesia

Luke Lucas*	London Metropolitan University
Yanbo Hu	London Metropolitan University

Is the sound "KIKI" sharp and spiky, or smooth and round? What shape does "BOUBA" take? The KIKI-BOUBA test has consistently shown that 80% of participants associate KIKI with something spiky, (Ćwiek, 2021) highlighting the phenomenon Cross-Modal Correspondence (CMC)—where sensory inputs from different modalities (e.g., sound and shape) create consistent pairings across a population. Synaesthesia, in contrast, occurs when individuals experience unique, personal cross-modal associations, such as seeing colours when hearing music. While both CMC and synaesthesia revolve around sensory perception and integration, CMC is widespread, and synaesthesia is rare, (Simner, et al., 2006).
This study explored intersections of CMC and synaesthesia by expanding both the variety of stimuli and the sensory modalities investigated. 20 audio and visual stimuli, were used, adding range to the sound-shape pairings. The incorporation of a third modality is uncommon in synaesthesia studies; emotions are an ever present feature of perception so this interoceptive modality was added. Past studies typically limited investigation to exteroceptive senses (e.g., sight, sound). However, perception is multimodal and not limited to exteroceptive modalities. The addition of interoceptive emotion addresses this limitation.

The hypothesis that the emotional response to a sound would correlate with the emotional response to the associated shape, and this would be more pronounced among the synaesthetes, was rejected after data analysis.
An unexpected correlation did emerge; a divergence in arousal levels between sounds and shape, significantly different between synaesthetes and non-synaesthetes. This alludes to modality differing, coexisting emotional interoceptive and exteroceptive processing.
This study offers fresh insights into the complex interplay between emotional and sensory processing, advocating for a multimodal approach to understanding synaesthesia, cross-modal correspondences ad auditory processing. By highlighting the role of emotion in sensory perception, it challenges conventional approaches and opens new avenues for exploring internal and external multisensory integration integration. .

Author e-mail: lul0200@my.londonmet.ac.uk

Back to Schedule

10:15 Naturalistic multimodal spatial interactions

Laurie Heller*	Carnegie Mellon University
Anjelica J. Ferguson	Carnegie Mellon University
Sungjoon Park	Carnegie Mellon University
Daniel Rosenberg	Carnegie Mellon University

We studied auditory-visual interactions under naturalistic conditions in which observers moved real objects in an open environment, thereby supplying cues from gesture and motion. With the use of a wavefield synthesis array, sound sources were invisibly placed to be either collocated with the object, or several degrees to the left or right of the object. With the use of motion tracking cameras, the sound sources could move along with the objects while observers moved them (in the congruent motion tracking condition). Because the sound was emitted from the speaker array and not the object, the sound sources could remain stationary while the object was moved (in the incongruent motion-tracking condition), and vice versa. We tested the hypothesis that auditory and visual motion congruence would promote the binding of sounds and objects, while incongruent motion would disrupt it. Our experiment began with a setup in which an object was stationary while a 200-ms 79 dB SPL broadband sound was played twice per trial, on opposite sides of the object. Twenty seated, stationary listeners indicated the direction of the change in sound location in a two-interval forced-choice task using an ipad interface. A three-down one-up adaptive track determined the lateralization threshold in degrees of azimuth. The threshold was higher when the object was present than when it was absent. This was an expected effect of the ventriloquist effect, in which observers tend to localize sound towards a nearby visual stimulus. In the congruent motion condition, when the sound was collocated with the moving object, lateralization thresholds were even higher. However, inconsistent with our predictions, thresholds were just as high in the incongruent motion condition. In contrast, conditions involving having objects on either side of the midline revealed lower thresholds.

Author e-mail: laurieheller@cmu.edu

Back to Schedule

10:30 Improving distance perception in virtual reality using spatially incongruent real-world sounds

Maggie McCracken*	University of Utah
Hunter Finney	University of Utah
Jeanine Stefanucci	University of Utah
Sarah Creem-Regehr	University of Utah

Virtual reality (VR) technology offers highly controlled immersive experiences that can be used to conduct spatial cognition research. However, there are VR-specific perceptual biases that reduce its ability to generalize all findings to real-world scenarios. One specific bias causes observers to perceive objects as closer than intended, a phenomenon called distance compression. Although prior research has focused primarily on visual and technological solutions to distance compression, using other senses as a potential remedy (e.g., hearing) for the bias has not been explored. Here, participants completed both a blind-walking task for distance estimation and a cross-modal perceptual-matching task to evaluate audiovisual spatial displacement. The results showed that real-world sounds spatially displaced both in front of and behind a corresponding visual target increased the accuracy of blind-walking, relative to vision-only cues, albeit with a small effect. Individual differences measures, including performance on the perceptual-matching task and predicted auditory weights, did not account for the variability in the effectiveness of spatially displaced sounds. Additionally, participants did not optimally integrate auditory and visual cues in a blind-walking task according to the Bayesian integration model. Overall, this study demonstrates the potential for adding multisensory information (specifically the addition of non-virtual sounds) to improve the accuracy of distance perception in VR.

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

Back to Schedule

10:45 Effects of meter-rhythm complexity and stimulus sound intensity on the vigor of synchronous walking

Gregory D. Shay* Retired Independent Scientist

Much research has focused on identifying the stimulating qualities and conditions of music that evoke 'groove'—the pleasurable human urge to move in time to music. However, no known studies have explored the potential groove effect of simple metronome meter-rhythms on repetitive physical activities, including that of walking. Additionally, while the effect of sound intensity of music has been examined during walking, no similar studies have included metronome stimuli. During walks in synchrony to the perceived beat over an outdoor 2-mile course, 120 bpm meter-rhythms created with an accenting metronome were compared with 120 bpm music for stimulating quality over a controlled range of sound intensity using an integrated single-case design. The order of increasing step length and proportional walking speed elicited by the four stimulus conditions was: nearly monotonic basic 10/4 metronome meter < basic 4/4 metronome meter < secondary accented 4/4 metronome meter < familiar–enjoyable 4/4 single-genre music—reflecting the order of decreasing meter monotonicity for the first two stimuli and the order of increasing rhythmic complexity overall. Although the sense of pleasure also increased with meter complexity, there was little awareness of the induced urge to move. Based on mean speeds for each metronome and music condition across the range of sound intensity, post-hoc Tukey HSD analysis revealed significant ($p < .05$) differences for five of six paired conditions. Over the studied intensity range of 59 dB (faint) to 87 dB (moderately loud), walking speed increased similarly (approximately 10%) for all stimuli with increasing intensity. This indicates that intensity serves as a positive modulator of the auditory stimulus. A power law was considered plausible for the intensity–speed relationship. Power regression fits of data were nearly linear over the intensity range. These findings have significant implications for rhythmic physical activities and may contribute to a deeper understanding of the human auditory sensorimotor system. Additionally, they may have potential for interventions in human gait disorders.

Author e-mail: shaygreg2828@gmail.com

Back to Schedule

Keynote (1:30-2:00)

Rhythms and Algorithms: From Vibrations in the Ear to Abstractions in the Brain

David Poeppel* New York University

The brain has rhythms - and so do music and speech. Recent research reveals that the temporal structure of speech and music and the temporal organization of various brain structures align in systematic ways. The role that brain rhythms play in perception and cognition is vigorously debated and continues to be elucidated through neurophysiological and computational studies of various types. I describe some intuitively simple but surprising results that illuminate the temporal structure of perceptual experience. From recognizing speech to building abstract mental structures, how the brain constructs and represents time reveals unexpected puzzles in the context of auditory perception and language comprehension.

Author e-mail: david.poeppel@nyu.edu

Back to Schedule

Auditory Scenes & Timing (2:00-3:00)

2:00 Did I do that? Investigating objective auditory-motor agency in younger and older adults

Alexis Basciano*	Arizona State University
Saraching Chao	Arizona State University
Alessandra Isaacson	Case Western Reserve University
Corianne Rogalsky	Arizona State University
Ayoub Daliri	Arizona State University

Agency, or an awareness that one is in control of their actions, is crucial to meet our movement goals and is thought to emerge from post-movement comparison between predicted and actual sensory outcomes. However, older adults and stroke survivors are known to experience changes in their sensorimotor systems due to age and injury respectively. For individuals with changes to their sensory subsystems, can cues during the planning stage of movement increase agency over their actions and subsequently improve movement outcomes? In a series of studies, older adults and stroke survivors completed Altered Auditory Feedback (AAF) paradigms. Here, their speech signals were altered in real time to observe changes in their speech in response to these artificial errors. Stroke survivors' adaptation and compensation to these artificial errors improved when sudden, but not gradual, perturbations were applied to their speech. Following the results of this study, younger and older adult groups completed an AAF compensation paradigm with premotor auditory probes while collecting ERP data to answer the following questions: Are there age related differences in auditory N1 sensory attenuation and compensatory speech behavior between young and older adults? If so, do prospective auditory probes presented during the planning stage of movement improve sensory attenuation and speech compensation? Data collection is ongoing, though N1 suppression and compensatory speech behavior are as expected. Older adults appear to have more variable speech responses. Following data collection, we plan to correlate the magnitude auditory N1 suppression with magnitude compensation for each subject to determine the relationship between the two and to detect differences between the age groups. Results from these studies have the potential to inform prominent models of speech motor control, and to guide future directions for motor speech rehabilitation for stroke survivors.

Author e-mail: abascian@asu.edu

Back to Schedule

2:15 Auditory perception of the material composition of everyday objects

Aiden Iveris*	University of California San Diego
Gedeon O. Deák	University of California San Diego
Sarah C. Creel	University of California San Diego

Environmental sound perception enables listeners to form rich conceptual representations which reflect salient features of their surroundings. Audition thus informs recognition of actions, events, and specific object qualities – including material composition. This study explores the specificity and stability of material representations elicited by perceiving naturalistic acoustic events recorded from everyday object interactions.

Auditory stimuli were recorded using a DPA binaural headset worn by an experimenter while they interacted with objects. Images of the same items, manipulated with the same hand shape and motion, were captured with a Pixel 8 Pro. Objects belonged to four categories (plate, cup, small bowl, and mixing bowl), each instantiated in six different materials: ceramic, glass, metal, wood, paper, and plastic. Sounds resulted from four interaction events (setting down, sliding, pouring water, and pouring cereal). This variety of actions and objects was used to ensure that performance reflected generalized material representations. Forty-eight adult participants listened to one sound at a time and chose between two images (one matching). Eye movements were also tracked to identify possible subsets of auditory features associated with decision making.

On trials where only object-material differed between response options, mean accuracy averaged 82% (SD = .03, $p < .0001$; for comparison, 99% on control trials where action and object-category also varied), indicating material perception was robust across a variety of contexts. Accuracy for specific materials varied from 74% correct (wood) to 88% (paper). Closer examination revealed differences in material confusability (e.g. glass vs. ceramic: 53%, but glass vs. paper: 95%, $p < .0001$). Hypothetically, highly confusable materials share more sound-influencing physical properties (weight, rigidity, surface texture), an explanation corroborated by consistent grouping of these materials during multidimensional scaling analysis on action and object-type subsets of the data. Ongoing analyses are examining eye tracking data, which may indicate diagnostic features in the acoustic stimuli.

Author e-mail: aiveris@ucsd.edu

Back to Schedule

2:30 Positive feedback, negative impact: The role of feedback in auditory temporal order judgment

Leah Fostick* Ariel University

Auditory temporal order judgment (TOJ) measures an individual's temporal resolution, typically involving two short sounds with a varying silent gap. Feedback is usually not provided in auditory TOJ tasks, which is unconventional because feedback can help participants identify mistakes, boost confidence, increase motivation, and enhance performance. Moreover, studies have shown that some participants struggle with spectral TOJ, a version involving tones of different frequencies, and their performance remains at chance level, even with an inter-stimulus interval (ISI) of 240 ms. This raises the question that maybe feedback can be beneficial in auditory TOJ, which was the aim of the present study.

A total of 304 participants were divided into eight groups and performed spectral or spatial TOJ using constant stimuli or adaptive procedures, with or without feedback. The stimuli were 15 ms pure tones of 1 and 1.8 kHz at 65 dB SPL. In the adaptive version, the ISI started at 240 ms and was adjusted based on performance, with the threshold calculated as the average of the last eight of 10 reversals. The constant stimuli version used eight ISI values between 5 and 240 ms, with the threshold set at 75% accuracy. Participants were categorized as High-level performers (High-LP) if their threshold was shorter than 5 ms, Mid-level performers (Mid-LP) with threshold of 5-240 ms, or Low-level performers (Low-LP) if failed to perform the task at 240 ms.

Feedback did not impact performance in spatial TOJ for either version, nor did it affect adaptive spectral TOJ. However, feedback in the constant stimuli version of spectral TOJ resulted in fewer Low-LP and more Mid-LP participants, with no change in High-LP rates. These results suggest that task characteristics, feedback type, or participant strategies may hinder feedback from improving performance, warranting further research into the underlying factors.

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

Back to Schedule

2:45 An empirical study of synchronization behaviour and experience in human-machine trios

Bavo Van Kerrebroeck*	McGill University
Marcelo M. Wanderley	McGill University
Alexander P. Demos	University of Illinois at Chicago
Caroline Palmer	McGill University

When musicians coordinate their actions to achieve successful synchrony, they rely on an intricate alignment of auditory and motor processes. Such alignment has traditionally been captured using either linear error-correction models or non-linear oscillator models, focusing on discrete timing or continuous and instantaneous phase adaptations respectively. While previous research has successfully used these models to reveal the principles and effects of synchrony in individuals and dyads, the dynamics of larger groups remain underexplored. Moreover, standard measures like pairwise asynchronies are inadequate for larger groups, as they do not scale effectively with increased group size. The current study evaluates synchronization and experienced social interaction in musical trios using an innovative virtual partner interaction paradigm. Forty-eight musically trained adults were organized into 24 trios and asked to perform musical synchronization tasks with a human confederate and with an adaptive virtual partner. Tone onsets of the virtual partner were generated according to one of three algorithmic models: linear error-correction, Kuramoto, or delay-coupled model. The trios produced simple in-phase and anti-phase isochronous rhythms. After each condition, the trio members rated their perceived social interaction.

Trios that contained a virtual partner exhibited significantly greater synchronization accuracy and stability (measured using clusterphase) compared to trios that contained a human confederate. Among the three adaptive models, trio performance with the delay-coupled model generated better synchronization stability than the error-correction or Kuramoto models; this finding held for the simple isochronous rhythms but not for the anti-phase rhythms. Finally, participants reported a higher perception of synchronization success, a greater sense of control, and stronger social connections when synchronizing with virtual partners compared to a human confederate.

These findings highlight the potential of virtual partners to not only enhance synchronization, but also the performers' subjective experiences. Thus, virtual partners could be valuable tools for improving group dynamics and performance quality.

Author e-mail: bavo.vankerrebroeck@mail.mcgill.ca

Back to Schedule

Speech & Language (3:15-4:30)

3:15 The Morgan Freeman effect: The advantages and disadvantages of voice familiarity when listening to speech while distracted

Manda Fischer* University of Western Ontario
Ingrid Johnsrude University of Western Ontario

We examined whether familiar voices are more efficiently processed (i.e., less resource intensive) than unfamiliar ones, by comparing their intelligibility when masked by a competing talker, while also manipulating concurrent cognitive load. Participants (N = 30, 17 Female and 13 Male) heard two sentences spoken concurrently in different voices (familiar-unfamiliar or unfamiliar-unfamiliar) and reported the content of one (target) while ignoring the other (masker). While listening to the target-masker pair, participants either tracked the location of four moving dots on a screen (dual task; DT) or ignored these (single task; ST). In both single- and dual-task conditions, target word-report accuracy was highest when the target voice was familiar. Accuracy was lowest when the masker voice was familiar; participants erroneously reported words from the familiar masker. The effect of DT on word-report depended on the familiarity of the target voice: word-report was higher in ST than DT for unfamiliar, but not familiar, targets. These results 1) indicate that familiar voices are more resource efficient than unfamiliar ones and 2) suggest that familiar voices enhance selective attention when task-relevant but distract when task-irrelevant.

Author e-mail: mfisch5@uwo.ca

Back to Schedule

3:30 Intrinsic voice memorability affects memory for spoken content

Cambria Revsine*	University of Chicago
Kristen Van Engen	Washington University, St. Louis
Wilma A. Bainbridge	University of Chicago

People tend to remember and forget the same faces, scenes, objects, and more, a result of their intrinsic stimulus memorability (Bainbridge et al., 2013; Isola et al., 2014). The memorability of auditory stimuli was characterized for the first time in a recent study, in which participants showed consistent memory performance for the voices of individuals when speech content was held constant (Revsine et al., 2024). However, it is still unknown whether memory for a speaker's voice affects memory for content—in other words, do people better remember what someone with an inherently memorable voice says? To test this, we used clips of speakers across a range of memorability, as determined in a prior experiment, but now each speaking a unique sentence. 25 participants performed a memory task in which, on each trial, participants heard a new sentence, performed a 2-back distractor task, and then repeated the sentence aloud from the beginning of the trial. Speaker and content were counterbalanced across two task versions. During the experiment, participants' pupils were tracked, as pupil size is known to be a proxy for cognitive effort (Pichora-Fuller et al., 2016). We found that the likelihood of making an error when recalling a sentence negatively correlated with the memorability (hit rate) of the speaker's voice. Pupil size was also positively correlated with recall error rate, and negatively correlated with speaker memorability, during and after the presentation of the stimulus and the recall period. These results suggest that speech by individuals with more memorable voices is better remembered. Furthermore, this relationship may be explained by memorable voices being easier to process, as they elicit smaller pupillary dilations. These findings have widespread implications whenever speech is remembered, such as in the classroom, on conference calls, when using virtual assistants, and more.

Author e-mail: crevsine@uchicago.edu

Back to Schedule

3:45 Contributions of distributed training and semi-supervised learning to improvement on a non-native phonetic classification task

Beverly A. Wright*	Northwestern University
Hannah R. Rostollan	Northwestern University
Ruijing Ning	Linköping University
Shendy Ng Cen	Northwestern University
Alexandra D. Murueta	Northwestern University

While analyzing data for control conditions in an investigation of learning on a non-native phonetic classification task, we noticed something surprising: participants who were initially tested 15 minutes after a single training session improved markedly more by the next day than those who were initially tested immediately after training. We interpret these outcomes in light of three known phenomena from the broader learning literature. (1) Distributed versus massed training: learning can be enhanced by spreading training over time (distributed training) rather than condensing it into a continuous training session (massed training). (2) Semi-supervised learning: learning that requires training with feedback can be induced or modified by combining periods of training with feedback and periods of training without feedback. (3) Testing effect: a close cousin of distributed training and semi-supervised learning, learning can be enhanced by including intermediate tests within the overall training regimen. The results point toward a simple means of improving the effectiveness of speech-perception training. The results also imply that speech learning ultimately engages the same basic mechanisms as other learning types--opening the door to a wealth of related knowledge that could further optimize speech-training regimens.

Author e-mail: b-wright@northwestern.edu

Back to Schedule

4:00 Decoding semantic categories for spoken words and images from EEG data

Emma Karn*	Fordham University
McCall E. Sarrett	Northwestern University
Abigail Laver	University of Pennsylvania
Albert Kim	University of Colorado
John C. Trueswell	University of Pennsylvania
Joseph C. Toscano	Villanova University

Previous brain signal-decoding work has shown it is possible to train a model to predict semantic categories (e.g., animals vs. tools) from MEG data for printed words or pictures. This demonstrates that distinct features corresponding to these different semantic categories can be detected at the scalp. Here, we ask whether abstract semantic representations can be decoded from scalp-recorded EEG data in response to spoken words and visual pictures during a semantic categorization task to examine both domain-specific and domain-general representations. We used a 2 (item category: animal vs. tool) x 2 (modality: spoken word vs. visual picture) repeated-measures design. Each semantic category consisted of 24 items, with three exemplars of each item. Each stimulus was presented twice, for a total of 576 fully randomized trials. Participants' task was to press a button indicating whether the stimulus corresponded to an animal or a tool (independent of whether it was an auditory or visual stimulus). EEG data were recorded and timelocked to the stimulus presentation. First, a support vector machine (SVM) classifier was trained to decode semantic categories within auditory and visual stimuli. Results suggested semantic category can be decoded above chance for both modalities. Visual decoding was particularly robust, with a peak around 100 ms after stimulus onset, which may correspond to low-level visual features, and a peak around 200 ms, which may reflect semantic differences. Auditory decoding was less robust but still above chance, with a peak around 200 ms. Second, we trained a classifier on cross-modal decoding (i.e., training on images and testing on spoken words, and vice versa). Decoding was less robust cross-modally, but showed potential later-occurring effects of semantic category decoding, suggesting the detection of a modality-general semantic representation. Overall, these results demonstrate robust decoding of semantic categories from EEG data for visual objects and spoken words.

Author e-mail: ekarn@fordham.edu

Back to Schedule

4:15 On the role of orthography and images in early foreign language learning: An ERP study

Mathew Cieśla*	Northumbria University; Adam Mickiewicz University
Efthymia C. Kapnoula	Basque Center on Cognition, Brain, and Language; Ikerbasque
Maksym Pozdniakov	Adam Mickiewicz University
Justyna Gruszecka	Adam Mickiewicz University
Katarzyna Jankowiak	Adam Mickiewicz University

Previous research on foreign language (LX) learning has largely focused on the use of lexical referents, or words, as the primary means of testing L2 development. However, it is well described that images result in better retention and recall of new items, as described in the picture superiority effect (PSE) and the bilingual/dual coding theories (b/DCT; Paivio & Desrochers, 1980; Paivio & Desrochers, 1980). The present study tested the effects of lexical and image referents on LX learning using artificial words. 30 Polish native speakers learned 40 LX words with either images or L1 (Polish) lexical referents, through a familiarisation phase which included auditory inputs of the LX words (Bakker, 2014). This familiarisation was followed by a four-alternative forced choice task (4AFC). Participants were then tested 24-hours later while event-related potentials (ERPs) were recorded, using a translational priming paradigm that included congruent and incongruent conditions to the modality in which the LX word was originally learned. The behavioural results showed a facilitatory effect of images. Specifically, LX words learned and tested with images resulted in higher accuracy rates and faster response times than words learned with lexical referents, which provides support to both the PSE and b/DCT. These findings were further corroborated by the ERP results, showing smaller Late Positive Complex (LPC) amplitudes for target works preceded by image primes, compared to lexical primes. Furthermore, larger N400 amplitudes were observed for congruent relative to incongruent learning-testing trials, indicating a more robust activation in the semantic network when retrieving words learned in the same modality. Together, these results suggest that images provide a more salient modality for L2 learning, in which visual referents generalise better to new vocabulary items than lexical referents. Also, word retrieval activates stronger memory representation when training and testing are conducted in the same modality.

Author e-mail: mat.ciesla@northumbria.ac.uk

Back to Schedule

Poster Abstracts

Speech & Language (1-10)

1 Engaging Stories for the Study of Attention and Audition: An introduction to the ESSAA database

Lauren Petley*	Clarkson University
Sarah Jane Gascoyne	Clarkson University
Claire McCarty	Clarkson University

The cognitive effort that individuals expend on difficult tasks is influenced by their motivation to perform well. Such motivation can be induced extrinsically using rewards or might be evoked intrinsically by using research protocols that are inherently interesting. Given recent trends towards using naturalistic stimuli to study hearing and attention, the present study characterizes a selection of short stories with respect to the engagement that they induce. Twenty-seven short written works were selected from the public domain whose spoken durations were between 10 and 20 minutes. Participants read the written stories, completed a 10-item factual quiz to verify attention, then completed Kuijpers' Story World Absorption Scale (SWAS). The SWAS is an 18-item instrument that measures reader engagement along four dimensions: transportation, emotional engagement, mental imagery, and attention, as well as producing a total score. A 30-second excerpt of narrative audio for each story, drawn from LibriVox, was also ranked for interest using a single Likert scale item. A total of 930 ratings were obtained from North American participants aged 16 - 25 via an online survey. Absorption scores varied considerably between some works and accuracy on the factual quiz tended to correlate with absorption scores (rho = 0.41, $p < .001$). Interest ratings for the narrators' reading styles, by contrast, rarely differed significantly. These results suggest that, while the choice of story is important for designing an interesting study, the choice of narrator is less relevant. Our results also stress caution in the use of factual quizzes to verify attention towards narrative stimuli that are not engaging. The ESSAA database comprises absorption and narrator interest scores, as well as the factual quiz items used for verifying attention. It is intended to help researchers design interesting studies and is freely available via the Open Science Framework.

Author e-mail: lpetley@clarkson.edu

Back to Poster Session

2 Intelligibility of medically related sentences in quiet, speech-shaped, and hospital noise

Sarah White*	University of Chicago
Alex Holly	University of Rochester
Tessa Bent	University of Indiana Bloomington
Erica Ryherd	University of Nebraska Lincoln
Melissa Baese-Berk	University of Chicago

Noise levels in healthcare settings often exceed recommended limits established by the World Health Organization. While previous studies have examined the impact of noise on speech perception in classrooms and workplaces—demonstrating decreased performance on working memory tasks and reduced intelligibility—the effect of noise on speech perception in medical environments remains underexplored. This gap in research is critical, as effective communication between healthcare providers and patients is essential for delivering quality care. Moreover, past research has typically focused on noise generated by buildings, equipment, or human activity. Although these sources are relevant to noise in healthcare settings, none have sufficiently addressed the impact of hospital noise on speech intelligibility. This raises an important question: how might hospital noise affect patients' ability to understand medical information? In this study, we examined speech perception under conditions of quiet, hospital noise, and speech-shaped noise through an intelligibility task with participants aged 65 and older. Our goal was to assess the effects of hospital noise on speech perception within a population more likely to interact with medical professionals. Additionally, we investigated the influence of word frequency and familiarity on speech intelligibility by utilizing both standard sentences and medically related sentences. We collected data on participants' hearing ability, cognitive acuity, and experience in hospital settings to explore other factors that might interact with speech perception in noisy environments. Preliminary results indicate that speech intelligibility is significantly better in quiet conditions compared to noisy ones, aligning with previous research. Standard sentences were also found to be generally easier to understand than medically related sentences across all listening conditions. Furthermore, likelihood ratio testing revealed that, in addition to the presence or absence of noise, vocabulary size and hearing ability are reliable predictors of speech intelligibility.

Author e-mail: sarsar10green@gmail.com

Back to Poster Session

3 Correlation between speech movement planning, task complexity, and beta frequency activity in normal adults

Mohsen Parsa* Arizona State University

Speech disorders are among the most common forms of communication disorders. Many speech disorders are associated with malfunctioning neural processes involved in speech production. Understanding these processes could enable us to develop new neuromodulation interventions for these disorders.

In the present study, we used electroencephalography to examine neural signatures of speech planning and production. Specifically, we examined the functional relevance of beta-band activity in speech planning and production. Studies of limb movements have shown that beta oscillation amplitude decreases before movement onset and increases after movement completion. Here, we aimed to determine whether the dynamic of beta activity during speech planning is sensitive to task complexity and planning time.
This study is ongoing; however, we have recruited ten healthy adults for our study. Each participant completed several blocks of speaking and silent reading tasks. Each block consisted of 48 trials. In each trial, a word in red appeared on the screen, and after a brief interval, the color of the word changed to green. In the speaking task, participants were instructed to say the word aloud. In the silent reading condition, participants were instructed to read the word silently (without movements and sounds). The silent reading task served as a control condition to control for the effect of visual processing of reading the target words. To experimentally manipulate task complexity, we used monosyllabic and multisyllabic target words. We used short (0.75 and 1.25 s) and long intervals between the red and green words to manipulate planning time.

We extracted the time course of beta activity throughout the task and used it as the dependent measure. Our preliminary results suggest that increased task complexity and longer planning times may lead to larger changes in beta activity before speech onset.

Author e-mail: mparsa1@asu.edu

Back to Poster Session

4 Equivalent processing of self-relevant information in autism: Evidence from the cocktail party phenomenon

Jan Philipp Röer* Witten/Herdecke University
Lejla Alikadic Witten/Herdecke University

The cocktail party phenomenon refers to the finding that approximately one third of the participants detect their own name in the irrelevant message during a selective listening task. Here we present a preregistered replication of the cocktail party phenomenon in autism. Previous studies on the processing of self-relevant information yielded mixed results. Some studies showed that the response to one's own name is diminished in autistic individuals, while others found no differences. However, a systematic investigation of the classic cocktail party phenomenon comparing autistic and nonautistic individuals has not yet been conducted. We used a standard selective listening task in which the task for the participants was to listen to the relevant message presented to their right ear and repeat each word as soon as they hear it while at the same time ignoring the irrelevant message presented to their left ear. The irrelevant message consisted of short sentences, two of which contained the participant's own name and that of a yoked-control partner, respectively. Replicating the classic finding, 13 of 48 participants in the autistic group (27 percent) and 18 of 50 participants in the neurotypical group (36 percent) reported hearing their own name in the irrelevant message and those who did made more shadowing errors shortly after the presentation of the name. There were no group differences and name detection was not related to individual differences in the autism-spectrum quotient.

Author e-mail: jan.roeer@uni-wh.de

Back to Poster Session

5 Navigating uncertainty: The influence of cognitive and non-cognitive factors on the interpretation of provisional language in science communication

Shruti Kate*	Kent State University
Katie Asaro	Kent State University
Jennifer M. Roche	Kent State University

When it comes to scientific communication, the use of provisional language (e.g., saying might), may play an important role in shaping interpretation. A pragmatic interpretation of a hedged statement may either aid or impede a listener's ability to effectively assess the veracity of a scientific claim. In the current study, we evaluate the ways in which cognitive (e.g., existing knowledge) and non-cognitive factors (e.g., believability, trustworthiness, and truthfulness) are impacted when scientific facts and myths are communicated with provision (i.e., hedged). We employed a computer mouse-tracking task to assess participants' evaluations of veridicality of provisionally modified spoken facts and myths. Results indicated that myths were rated as generally less believable, trustworthy, and seemed less true than facts. However, individuals with lower germ theory literacy were less impacted by the presence of a hedge, relative to the individuals with higher germ theory literacy. The findings indicated that a hedge may have acted as a pragmatic cue for non-literal interpretation of a statement for the high literacy group, but the low group did not make this differentiation. In fact, the evaluation of the action dynamics of responding showed that the more accurate knowledge a participant had about how germs spread, the more cognitive competition they exhibited when assessing a hedged falsehood — likely because existing knowledge competed with the acceptance of the assertion, as the participant attempted to infer why someone would make a falsehood sound provisional. Provisional language used in spoken communication about science may cue some listeners that a pragmatic interpretation is warranted, while serving no purpose for other listeners. Therefore, hedging may promote the use of pragmatics to guide a listener in easy and fast interpretation, but as communicators of science, we should consider that knowledge will interact with interpretation of provisional language in important ways.

Author e-mail: skate@kent.edu

Back to Poster Session

6 A meta-analytic perspective on characteristics of brief speech alerts

Michael A. Nees* Lafayette College

Brief auditory alerts such as earcons, auditory icons, spearcons, and brief speech messages are common forms of auditory displays for providing status alerts in practical applications. A previous systematic review and meta-analysis of brief auditory alerts (Nees & Liebman, 2023, https://doi.org/10.1080/25742442.2023.2219201) showed a pattern such that speech and speech-based alerts were generally superior to auditory icons and earcons with respect to reaction time and accuracy of performance, with speech and speech-based alerts also superior to earcons with respect to subjective ratings. Given the apparent promise of speech alerts, the current follow-up analysis re-examines speech alerts in the same data set to look for moderating effects of: (1) the gender of the voice used in speech alerts; and (2) the use of naturalistic speech recordings versus synthetic/text-to-speech (TTS) recordings for speech alerts. An unweighted random effects model is used to examine the potential effects of gender and naturalistic/TTS speech on reaction time, accuracy, and subjective ratings outcomes. Results are discussed in terms of potential implications for auditory display design and future research, as well as limitations of the analyses presented.

Author e-mail: neesm@lafayette.edu

Back to Poster Session

7 Assessing AI generated speech: The impact of gender, (in)directness, and intonation on the perception of rudeness

Jacob A. Miller* Kent State University
Jennifer M. Roche Kent State University

With advancements in text-to-speech technology, AI may offer control and faster stimulus creation of naturally sounding acoustic stimuli. To determine if AI generated speech is a useful tool to create vocal speech stimuli, we used elevenlab's speech generator to create three types of spoken statements: direct (Go to the store.), indirect (Can you go to the store?), and neutral (This is a store.) and asked participants to provide social judgments of rudeness in a computer mouse-tracking task. Direct communication is often perceived as more rude than indirect, especially when used by women. The study's stimuli included relatively flat intonation contours for the AI woman's productions, but the AI man's voice was marked by declining intonation. To control for talker characteristics that could bias a listener's interpretation of rudeness, we digitally imposed a flattened pitch contour on all sound files. Results indicated that direct statements were more rude than indirect and neutral statements. The AI woman was explicitly rated as more polite than the AI man, even when she produced direct speech — potentially reflecting social desirability bias. In fact, when evaluating the participant action patterns, the implicit processing path reflected a trending increase in cognitive competition when judging the woman's direct speech, aligning with the literature that finds women should be stereotypically expected to follow prescribed communication styles (e.g., be polite and agreeable). Nevertheless, we wanted to determine whether or not the effect may have been due to the extra-linguistic cues the AI generator produced. We then evaluated the acoustically modified vocal productions (i.e., flat intonation contour), the AI woman was no longer perceived as more polite than the AI man. The findings suggest it is important for researchers to evaluate and control for extra-linguistic parameters produced by AI generators.

Author e-mail: jmill365@kent.edu

Back to Poster Session

8 Physical exertion as a novel, in-the-moment index of listening effort

Carson Rumble-Tricker*	University of Guelph
Gurjit Singh	University of Toronto; Toronto Metropolitan University
Mark Fenske	University of Guelph

Attending to and understanding speech in noisy environments is challenging. Measuring corresponding increases in listening effort is important for investigating the cognitive mechanisms involved in speech perception, and for assessing the effectiveness of interventions that address hearing impairments. We're exploring the usefulness of physical exertion as a novel index of listening effort, having discovered that participants will exert physical effort (repeated button-pressing) to secure easier listening conditions (less background noise) in a subsequent speech-identification task. Levels of exertion matched task demands across differences in sentence predictability and initial background noise, suggesting that key-pressing allowed participants to avoid the need to exert listening-related cognitive effort.

We previously measured physical exertion before each trial. Here we ask whether button-pressing behaviour can also provide an in-the-moment index of changes in listening effort. Spoken-narrative stories were presented with gradually-increasing levels of background noise. Participants could reduce the noise, if desired, by pressing a spacebar. A progressive-ratio schedule, however, meant that increasing numbers of key-presses were needed to obtain easier listening conditions during each of six successive portions of a given narrative. Moreover, the order of these narratives were either left intact (coherent story) or scrambled (incoherent story). Participants' key-presses significantly increased as the progressive-ratio required more physical-exertion to obtain easier listening conditions. Self-reports after each narrative revealed that the scrambled-incoherent stories elicited greater subjective effort and boredom, and less attention and absorption, than the intact-coherent stories. Multiple-choice test performance likewise revealed better comprehension of and memory for intact-coherent stories than scrambled-incoherent stories.

Together, our findings provide converging support that understanding speech in noise is effortful, and that the costs of such cognitive effort are readily exchanged for increases in physical effort. Overall, the current results thereby suggest that measures of physical exertion can provide a path towards a useful in-the-moment index of listening effort.

Author e-mail: crumblet@uoguelph.ca

Back to Poster Session

9 English-Spanish bilinguals' perception of ambiguous speech sounds changes based on task language

Gabriella Thomas*	University of San Diego
Camille Torre	University of San Diego
Denise Picazo	University of San Diego
Laura Getz	University of San Diego

Speech perception is surprisingly complex, especially for bilingual speakers, who need to know the rules for phoneme pronunciation in multiple languages. One acoustic cue for phonemes is voice onset time (VOT), a measure for the length of different stop consonants. In English, voiced stop consonants like /b/ have short VOTs (around 0ms) and voiceless stop consonants like /p/ have longer VOTs (around 40ms). In Spanish, the same sounds are shifted in VOT, such that /b/ is pre-voiced with a VOT around -40ms and /p/ has a VOT around 0ms. Thus, an English voiced phoneme and a Spanish voiceless phoneme have identical VOTs. Our research investigated how bilingual English-Spanish speakers shift their perceived VOT boundary based on language context. Researchers interacted with participants in English or Spanish, after which participants completed an experiment where they were asked in a two-alternative forced choice task what they perceived for a continuum (-40ms to +40ms in 10ms intervals) of VOTs using three word/non-word pairs in English and/or Spanish (e.g., "bolo/polo" where both are words in Spanish but only "polo" is a word in English). We found VOT perception differences between participants completing the task in English and Spanish, especially at intermediate VOTs. For example, with the word pair "bolo/polo", bilinguals in the English group exhibited a rapid switch in categorization around -20ms VOT, consistent with the perception of the word over non-word endpoint. In contrast, bilinguals in the Spanish task reported hearing "bolo" more frequently until 0ms, consistent with the voiceless boundary in Spanish. Similar trends were observed for the basta/pasta and belly/pelí continuums, although the results were not as straightforward as anticipated. Our next step is to employ EEG/ERP analysis to investigate the time-course of how participants perceive sounds differently based on the task language.

Author e-mail: gabriellathomas@sandiego.edu

Back to Poster Session

10 Difficulty of speech-in-speech recognition: How does listening environment modulate informational masking?

Madelene Tavarez*	Penn State University
Andrea Burgos Mercado	Penn State University
Navin Viswanathan	Penn State University
Annie Olmstead	Penn State University

Speech-in-speech recognition is a challenge that listeners often encounter. Masking in these situations occurs due to a combination of energetic and informational masking. Energetic masking arises from the physical overlap between the target signal and the masker, while informational masking arises from cognitive, attentional, and other factors. The contributions of informational masking are poorly understood. In Experiment 1, we manipulated informational masking by varying the target-masker onset asynchrony while holding energetic masking constant. To accomplish this, we altered the onset of the target without altering the overlap portion of the babble and target. We did so in the context of a phenomenon called Linguistic Release from masking—the observation that the intelligibility improves if the masking language is different from that of the target. Monolingual English listeners transcribed BKB sentences by a female talker with an English or Mandarin two-talker babble as maskers. The targets started either simultaneously with, 500 ms after, or 1000 ms after the masker. Listening accuracy was higher for the Mandarin than for the English masker, replicating typical LRM. However, the size of LRM was moderated by onset asynchrony with larger effects for longer lags. Specifically, performance with the English masker worsened with increased lag demonstrating that the detrimental effect of the English masker accrues. These results were obtained when the sex of the target and the masker were matched. An open question is whether a similar pattern will obtain if the target and masker are mismatched, making energetic masking weaker. In a follow-up study, we investigate the effect of lag on LRM when the target and the masker are mismatched on sex. Monolingual English listeners will transcribe the same sentences as in Experiment 1 but with male maskers. The results of this study will indicate the generalizability of lag effects under different conditions of energetic masking.

Author e-mail: mjt6180@psu.edu

Back to Poster Session

Music (11-18)

11 From performer to listener: Understanding emotional perception in music

Emma Ning*	University of Illinois Chicago
Claudia Zaragoza	University of Illinois Chicago
Jenna Hickey	University of Illinois Chicago
Roshni Jacob	University of Illinois Chicago
Alexander P. Demos	University of Illinois Chicago

We listen to music not only to maintain cultural and social identities but also to modulate emotions. Previous research on musical emotions typically falls into four theoretical frameworks: the discrete emotion model, the component process model, the dimensional models, and the postmodernism model. However, much of this work focuses mainly on the listener's perspective. Our research addresses this gap by acknowledging the bidirectional nature of music, involving both performers and listeners, and examining how listeners perceive musical emotions as experienced by the performer making the music. We conducted a study measuring the skin conductance responses (GSR) of a pianist performing the same musical piece in two styles—expressive and deadpan—and compared these with the GSR responses of the listeners. A pianist with 10 years of experience played Mozart's Fantasia in D Minor (K.397) on a Yamaha P125, with a Shimmer wireless GSR strapped to her left ankle. Both performances lasted 2 minutes and were played at the same tempo. The deadpan style minimized dynamic variations, while the expressive style incorporated them as typically desired in musical performances. The listeners (n = 92) were then asked to listen to the two styles, with GSR recorded from the same location as the performer. Using recurrence quantification analysis, we found significantly higher overlap between the performer and listener GSR responses for both styles compared to chance, falsifying the postmodernism model. These findings suggest that music conveys information from the performer to the listener.

Author e-mail: yning7@uic.edu

Back to Poster Session

12 Statistical learning: How contingent regularities influence learning, memory and preference for melodic pairs over the time-course of learning

Pierce C. Johnson*	University at Albany, SUNY
Gregory E. Cox	University at Albany, SUNY
Ron Friedman	University at Albany, SUNY

We used novel musical material to investigate how perceptual information is transformed into memory representations that enable prediction, and how the acquisition of such representations is related to motivational factors like preference. Using memory to form expectations is thought to be critical to the aesthetic experience of music, which may in turn reflect a general tendency to experience reward for making correct predictions (Huron, 2006). In two experiments, we addressed three questions concerning relationships between memory, prediction, and preference: How do we learn contingencies within a continuous stream of information? Do contingencies shape preference differently for information with different predictive value? And how does memory develop when predictive information is differentially informative about what one should expect to happen next? In the first experiment, participants listened to a continuous stream containing a mixture of single melodies and contingent melody pairs (e.g., melody A was always followed by melody B). After exposure, participants preferred contingent melodies over novel melodies and correct memory for pairs was correlated with preference for melodies that confirmed (melody B) rather than enabled (melody A) predictions, consistent with the hypothesized relationship between memory and preference for having predictions confirmed. In a second experiment, currently underway, the exposure stream contained pairs with varying levels of contingency, from 100% (B always follows A) to 50% (B follows A half the time). This second experiment will shed light on the time-course over which contingencies are learned; we predict faster learning of stronger contingencies. Experiment 2 also tests the hypothesis that there is an inverted-U relationship between contingency and preference, where moderate levels of contingency are preferred over strong or weak levels of contingency (Berlyne, 1971).

Author e-mail: pjohnson4@albany.edu

Back to Poster Session

13 Exploring the impact of auditory manipulations on piano learning: Sensory deprivation, imagery, and skill transfer

Renan Moreira Madeira*	University of Illinois Chicago; Federal University of Rio Grande do Sul
Regina Antunes Teixeira dos Santos	Federal University of Rio Grande do Sul

Several studies have investigated how the manipulation of auditory information affects temporal and expressive parameters of music performance. However, to what extent these effects translate into more ecological musical practice settings remains largely unknown. We designed an experiment in which four piano students learned piano works by Franz Schubert (Ländler D.145 no.3, Écossaisse D.781 no.5, Écossaisse D.781 no.4, and Waltz D.365 no.6) while having auditory and kinesthetic information manipulated. Four conditions were employed: (a) practice without the piano, similar to mental practice; (b) practice using only a recording of the stimulus piece, without the piano; (c) practice on a turned-off piano; (d) auditory learning, with the recording of the stimulus and the piano. A second practice session consisted of the practice of the same pieces in conventional conditions (with the piano and the score). Participants had their practice sessions videotaped, and the analysis consisted of the quantification of observable behaviors and their relations with the participants' goals during practice. While all participants resorted to some form of aural imagery during the silent practice conditions (a & c), only three of them managed to create accurate images (in terms of melodic contour), based on their aural skills when manipulating the stimulus pieces. Also, the conditions of deprivation that were employed seemed to have raised the participants' awareness to extract information from distinct sources of information (sound only or score only) to approach the practiced materials. These results suggested that while sensory restrictions during practice may enhance a musician's awareness of distinct information sources, the successful transfer of skills to conventional practice depends on the degree of structural similarity between the practice conditions and the actual stimulus.

Author e-mail: rmore24@uic.edu

Back to Poster Session

14 The Musician Recognition Task: A new measure assessing musical exposure as it relates to musicality

Veronica Kandro* University of Maryland
L. Robert Slevc University of Maryland

What defines a person as being more musical than another? Research assessing individual differences in musicality typically relies on self-reported formal training or performance on musically-relevant tasks (e.g., listening for differences between musical sequences). However, there are many components to musicality. Could exposure to music potentially play a role? Here, we take inspiration from evidence that exposure to literature (assessed with the "Author Recognition Task") is an important predictor of language skills to develop a similar measure of exposure to music. In this "Musician Recognition Task (MRT)," participants select names they recognize as musical artists from approximately 130 names, of which half are real musical artists and half are foil names. The artist list was created by selecting 400 artists (50 from each of eight genres) from Billboard's "Greatest of All Time" charts, cross-referenced with charts from other online music publications: AOTY, Classical Music Only, and Forbes. Foil names were created using generative AI. From this sample of 800 names, we selected a subset of musical artists ensuring a range of popularity (for our target college-aged population) using Spotify's API popularity score (which assigns an artist a numerical score between 0-100, taking into account the artists' streams, engagement, etc.). The MRT was administered alongside widely used measures of musicality (the Gold-MSI and PROMS), measures of genre preferences, personality characteristics, sensitivity to musical reward, and demographic information. Preliminary data show that participants performed significantly above chance on the MRT and that performance was highly variable. MRT scores correlated highly with subscales of the GoldMSI (particularly the active engagement and emotions subscales) and significantly, but less strongly, with performance on the PROMS. Together this suggests the MRT is a novel way to measure exposure to music and that musical exposure itself might play a unique role in musicality.

Author e-mail: vkandro@terpmail.umd.edu

Back to Poster Session

15 Melodies are not a useful mnemonic cue for immediate word recall regardless of familiarity

Isabella Lea Ramirez*	University of San Diego
Lauren O'Donnell	University of San Diego
Laura Getz	University of San Diego

Mnemonic devices and jingles paired with a musical melody have long been used as memory aids. Though this perceived phenomenon has been recognized in previous studies, research is mixed on the effectiveness of utilizing music to improve learning and memory. Furthermore, previous studies have not addressed melody familiarity as an aid for recall/recognition. Our experiments tested the effects of unfamiliar and familiar melodies on the memorization of unrelated word lists compared to a spoken control condition. Participants heard word lists with two-syllable words of mixed positive, negative, and neutral valence 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 complete an old-new recognition task. In two experiments, we found spoken text to outperform (Experiment 1 recall and recognition, Experiment 2 recognition) or equal (Experiment 2 recall) the various sung text conditions. These results indicate that regardless of familiarity, melodies may be distracting on initial presentation, due to working-memory capacity limits. Our initial results suggest that the colloquial belief of music aiding memory does not hold true for short-term recall of unrelated word lists. To understand potential factors influencing music as a memory aid, Experiment 3 will test the effect of melodic characteristics on recognition; we will compare participants who receive a single familiar melody to those who receive multiple melodies to learn a word list. Experiment 4 will examine the choice of words used for the lists; groups will learn an unrelated word list, a related word list, or a sentence with a familiar melody to determine which combination of words works best for memorization. These results will help identify what specific characteristics assist or inhibit music as a mnemonic device.

Author e-mail: isabellaramirez@sandiego.edu

Back to Poster Session

16 Is adolescence a sensitive period for acquisition of musical knowledge? Behavioral evidence in the light of pop music style change analysis

Annabel J. Cohen*	University of Prince Edward Island
Kristen Gallant	University of Prince Edward Island
Corey Collett	University of Prince Edward Island

There is a longstanding debate regarding sensitive periods for language acquisition, generally focused on childhood years. It has been proposed that adolescence may be a sensitive period for acquiring social knowledge (Blakemore & Mills, 2014). The prominence of music in the social lives of adolescents suggests biological significance of an adolescent sensitive period for music. The present research tests a theory that proposes the establishment of a mentally represented music grammar in adolescence that offers lifelong privilege for encoding of music consistent with that grammar, in contrast to music that violates the grammar, such as music in earlier or later styles. This Plasticity Theory of Implicit Music Knowledge Acquisition (PTIMKA) assumes that the grammar of popular music changes across decades. Recent support for this assumption comes from Hamilton and Pearce (2024) who analyzed the Billboard top five songs from the 1950's onward (Billboard Melodic Music Dataset, BiMMuDa), revealing systematic changes by decade on eight music-structural dimensions. We applied these analyses to the music hits used in our studies and found similar patterns. For example, across six decades, note onset density significantly increased while pitch standard deviation and interval size pairwise variability (nPVI) decreased. Our study compared knowledge and short-term retention of brief excerpts of top hits by two groups of participants differing in age (n = 19, 20.3 [2.3] years and n =26, 65.5 [6.7] years). Familiarity, knowledge of title and artist, judged year of popularity and short-term retention of popular music across six decades correlated significantly with several of the music-structural BiMMuDa variables: the sign of the correlation dependent on participant age (due to high behavioral scores associated with songs popular during adolescent years). These data offer support for an adolescent sensitive period for acquiring knowledge of grammar of music to which one is exposed to at that time.

Author e-mail: acohen@upei.ca

Back to Poster Session

17 Developing written music literacy: The relationship between working memory and melodic dictation ability

Erica R. Knowles*	Berklee College of Music
Leslie Anne Harrison	Berklee College of Music
Ethan Bessette	Berklee College of Music
Mi-Lan Hoang	Berklee College of Music
Rose Koo	Berklee College of Music
Bathanie Liu	Berklee College of Music
Ella Shalev	Berklee College of Music

Dictation, the ability to translate aurally presented musical phrases into music notation, is a major component of undergraduate music training. However, there is a paucity of research into the skills required to be successful in dictation, and the best pedagogical approaches to supporting student learning. This study aimed to consider the relationship between working memory (WM) and dictation ability in students early in their written music literacy development. Additionally, we were curious about the strategies students implement in the face of increasing WM load, as they may have pedagogical implications. Participants were 25 first-semester undergraduates placed in an introductory ear-training course. At pre-test, participants completed a standard measure of WM (digit span (DS)) and a measure of musical WM (Musical Ear Test (MET)). The same WM tasks were completed again after 8 weeks of instruction, alongside a music dictation task. Participants completed three dictations of increasing complexity to consider the impact of WM load. Dictations were scored on accuracy (melodic, rhythmic, and overall) and schema-use (interval pattern accuracy). While there was a significant correlation between scores on DS and MET, DS did not correlate with any of the music dictation scores, nor was there a significant change on any WM scores from pre-test to post-test. For the least complex dictation, higher MET scores were related to greater overall dictation accuracy. Interestingly, as dictation complexity increased, students relied on schema-based strategies to a greater extent, and MET scores (melodic subtest) correlated with interval pattern accuracy. Our results suggest that melodic dictation is related to domain-specific WM ability, and that schema-based strategies may play a role as WM load increases. Pedagogical techniques that develop skills related to WM and musical schema knowledge may reduce WM load and increase retention, leading to a better chance of successfully translating aural information into written notation.

Author e-mail: eknowles@berklee.edu

Back to Poster Session

18 Absolute Pitch among non-global possessors: Timbre, categorization and musical pleasure

Ébano Resende de Souza* Austrian Academy of Sciences

Perfect or absolute pitch (AP) is typically considered to be an asset for any musician, however it is not well understood and may have some drawbacks. From a case study among students from the Musicology and Analysis Department of the Conservatoire National de Musique et Danse de Paris - CNSMDP (under direction of Dr. Adrien Mamou-Mani - CNSMDP/IRCAM and Dr. Michèle Castellengo - LAM), we noted important issues raised by the students: 1) confusion and misunderstanding about the AP definition, 2) difficulty among students in identifying whether or not they have AP, 3) the relevance of timbre for AP identification among these students and, 4) students that had partial AP abilities reported difficulties in some musical activities. Among our findings we can highlight: 1) the strong relationship between AP identification and primary instrument; 2) the importance of the piano (pointed as the easiest timbre for recognition) and the voice (which occupies an important place among the hardest timbres for recognition, due, probably, to the overlap between AP label and the lyrics of the song) during AP identification and, 3) the loss of musical pleasure indicated by a considerable number of AP possessors. This last highlight especially calls for further research. We are therefore currently developing a project to understand to what extent AP perception can hinder RP perception and therefore the joy of music.

Author e-mail: ebano_ctu@hotmail.com

Back to Poster Session

19 Impact of sound source motion on the perception of occlusion and disocclusion.

Mike Russell* Bellevue University

Real-world settings typically contain numerous objects. In certain circumstances (e.g., observer location, position and size of the objects), one object may partially or fully occlude another object. Despite the common existence of clutter in everyday settings, relatively few studies have investigated the impact of clutter on auditory perception or have examined the ability of individuals to detect occlusion when relying solely on sound. Previous research, for example, suggests individuals are fairly capable of determining whether a sound source was occluded. However, that research involved stationary observers and stationary sound sources. As can be imagined, a change in the point of observation and/or the location of a sound source could result in a sound source becoming occluded or unoccluded. It can also be imagined that altering the location of a sound-producing object or the point of observation could influence the ability of observers to determine if an object is occluded. In the present study, participants listened to audio recordings of a sound source. The recordings varied in terms of the degree of occlusion. The task of the participants was twofold. First, participants indicated the degree to which a particular event involved any occlusion at all. Second, participants indicated whether the sound-producing object was occluded at the end of the recording. The findings of the present study enhance our understanding of how we perceive the world by sound and how motion influences that perception. The findings will also be considered with respect to the phenomenon of representational momentum.

Author e-mail: mirussell@bellevue.edu

Back to Poster Session

20 Context-dependent auditory inference in environmental sound recognition

Keland Moore* University of Iowa
James Traer University of Iowa

In everyday listening, humans transform continuous sounds into multiple perceptual categories (e.g., what object made that sound? Was it heavy? Hard? Metallic?). Such judgments are robust across different listening contexts (e.g., different object motions, orientations, etc.). Using listening tasks and Ideal Observer Models, we investigated which acoustic cues underlie successful auditory physical inference in variable (i.e., unpredictable) contexts. Specifically, we asked listeners to identify multiple attributes of objects (shape, weight, and material) from clattering sounds. Prior studies of speech perception have highlighted the importance of context-dependent inference (i.e., joint-inference, or the use of different acoustic cues in different contexts, such as different speakers, accents, etc.), and we investigated its role in this task. Our stimuli varied across all physical attributes and thus, for any query (e.g., weight) the other two attributes (shape, material) served as unknown and variable contexts. Additionally, we presented stimuli in two "acoustic contexts": (1) single impacts; and (2) naturalistic clattering sounds with bounces and scrapes. Models show different acoustic cues best predict human judgments in each task and condition, suggestive of context-dependent inference. The use of context-dependent inference in non-speech categorization suggests these processing mechanisms are not specific to speech perception but instead may be fundamental to hearing.

Author e-mail: keland-moore@uiowa.edu

Back to Poster Session

21 Pupil-linked arousal tracks adaptive auditory belief updating in spatially and temporally dynamic environments

Roman Fleischmann*	Austrian Academy of Sciences
David Meijer	Austrian Academy of Sciences
Ulrich Pomper	University of Vienna
Michelle Spierings	University of Vienna
Robert Baumgartner	Austrian Academy of Sciences

Auditory perception is subject to sensory noise and rapidly changing environments. To deal with ambiguous input, the auditory system needs to find the correct balance between flexibility and robustness, integrating sensory input with prior beliefs. Bayesian perceptual inference determines the statistically optimal solution. The locus coeruleus (LC) arousal system is suggested to mediate belief updating by amplifying sensory input through the release of noradrenaline (NA) at the cost of prior beliefs. We presented 49 participants with auditory sequences of random lengths, inducing either temporal movement (acceleration or deceleration, via manipulation of stimulus onset asynchronies, (SOAs)) or spatial movement (clockwise or counterclockwise motion, via manipulation of the stimulus's horizontal localization), while keeping the other dimension constant. Participants were tasked with discriminating the last direction of spatial or temporal movement in a two-alternative forced-choice design. We recorded task-evoked pupil sizes via pupillometry as a proxy for LC-NA activation. Sporadic change-points (CPs) within the sequences forced strong belief updating to maintain precise perception in the face of change. To increase ecological validity compared to common CP paradigms, we varied the level of sensory evidence continuously and at a rapid pace to elicit a broad spectrum of continuous, fast, online belief updates. We designed the experiment as a low-level perception task, asking participants only for discrimination, not for explicit predictions. We further chose stimuli with real-life connotations, emulating horizontal movement and tempo changes. A Bayesian CP model was fitted to the behavioral responses to estimate momentary surprisal, a precision-weighted quantification of the prediction error. We show surprisal to predict task-evoked pupil size on a rapid stimulus-to-stimulus level. Importantly, the relationship between surprisal and pupil size proved to be independent of the tested condition. The results support the notion of a modulation of auditory belief updating by the LC-NA arousal system, independently of perceptual domain.

Author e-mail: roman.fleischmann@oeaw.ac.at

Back to Poster Session

22 Investigating cognitive disruptions in misophonia: The impact of trigger sounds on working memory performance

Michael A. Tollefsrud* Kansas State University
Matthew G. Wisniewski Kansas State University

Misophonia is a condition characterized by strong emotional and physiological responses to specific sounds, often referred to as "triggers." These responses are often self-reported to interfere with ongoing cognitive processing. This project aimed to measure the cognitive disruptions experienced by individuals with misophonia using a behavioral assessment of working memory. Participants were recruited and scored on the Duke Misophonia Questionnaire (DMQ) impairment subscale. Participants completed a tonal pitch-matching task in which a 2-second retention interval was filled by misophonia triggers (e.g., eating cereal), neutral sounds (e.g., sweeping chips with a broom), or silence. We predicted that the Misophonic group would show significantly impaired pitch-matching accuracy when triggers were presented in the retention interval compared to neutral stimuli or silence. We also predicted that the control group would show similar performance between trials presenting trigger and neutral sounds during retention. Though there were minimal interference effects at low DMQ scores, high DMQ scores were associated with worse performance for trigger and neutral sound trials than silence. Implications, possible explanations, and directions for future work are discussed.

Author e-mail: matollef@ksu.edu

Back to Poster Session

23 Non-auditory impacts of wearing hearing protection: concurrent sensorimotor tracking and choice to wear

Matthew G. Wisniewski*	Kansas State University
Patty McClain	Kansas State University
Chelsea N. Joyner	Kansas State University

Occupational hearing loss (HL) is a significant problem despite employer encouragement to wear hearing protection devices (HPDs). Performance while wearing HPDs is one self-reported reason why workers choose not to wear them. However, few studies have supplemented these subjective reports with objective measures. Where they do exist, assessed performance measures have mostly characterized auditory situational awareness in gross terms (e.g., average speech comprehension scores over an entire session). Objective measurement of HPD impacts on non-auditory aspects of work performance and choice-to-wear are largely unknown. We characterized both in this study. Listeners heard stimuli sourced from the Coordinate Response Measure (CRM) corpus (i.e., sentences of the form "Ready <call sign> go to <color> <number> now). These commands informed listeners of which of nine moving on-screen objects to track with a computer mouse (e.g., "blue four" = blue square). The commands were presented in background street noise and heard under either No HPD or Simulated HPD conditions. Continuous recording of tracking error allowed examination of how HPD-wearing impacted speech comprehension and the accuracy of tracking. To characterize choice-to-wear, we asked listeners to choose between wearing an HPD or not wearing an HPD with different balances of reward (e.g., 2x reward for HPD vs. 1x reward for unprotected). Listeners spent less time tracking the correct object in the HPD condition. After trimming tracking data to those time points in which the target object was known, worse tracking accuracy was still exhibited by the simulated HPD condition. Further, listeners needed more reward to "wear" an HPD. The difficulty of the tracking task shifted this bias even further. Workers' complaints of poorer performance while wearing HPDs are justified and extend beyond just diminished auditory situational awareness. Addressing these issues will be an important part of addressing HPD non-use in occupational settings.

Author e-mail: mgwisniewski@ksu.edu

Back to Poster Session

24 Test of a selective entrainment hypothesis using concurrent tone sequences

Toni Smith* Michigan State University
J. Devin McAuley Michigan State University

The selective entrainment hypothesis put forward by McAuley et al. (2020; 2021) proposes that selective attention to speech in difficult listening situations is facilitated by entrainment by the rhythm of target speech but can also be disrupted by inadvertent entrainment by the rhythm of to-be-ignored background speech. Two predictions follow: (1) speech understanding in noise will be compromised when the natural rhythm of target speech is disrupted, and (2) speech understanding in noise will improve if the rhythm of a competing background is disrupted. A series of experiments tested and confirmed these two predictions by altering the natural rhythm of target speech and background speech (McAuley et al., 2020; 2021; Smith et al., 2024). The purpose of the present study is to extend this work on selective entrainment to non-speech auditory stimuli, focusing on the background rhythm effect. Participants listened to isochronous sequences of tones and were asked to judge whether the interval between a final pair of tones was shorter, the same as, or longer than the intervals between the preceding tones. These sequences were either presented alone or with a concurrent tone sequence of a different pitch, which participants were instructed to ignore. The to-be-ignored sequences were either isochronous or temporally irregular. Based on the selective entrainment hypothesis and prior work showing that duration judgments are sensitive to rhythmic context (e.g. McAuley & Jones, 2003), we predict that temporal judgments about a target stimulus (1) will be less accurate in the presence of a competing rhythm and (2) will be biased in the direction of the competing rhythmic context, (3) but only when the competing rhythm is temporally regular. Results will be discussed in the context of the selective entrainment hypothesis and Dynamic Attending Theory.

Author e-mail: smit3094@msu.edu

Back to Poster Session

25 Perceptual discrimination of changes in bit-depth resolution

Benjamin H. Stone James Madison University
Michael D. Hall* James Madison University

Individuals can discern highest- from lower-quality digital audio, but existing evidence tends to confound sampling rate and bit depth (see Mizumachi et al., 2014). The present study aimed to evaluate bit-depth-related perceptual ability using brief, complex tones to determine where increases in bit-depth resolution are no longer perceptually beneficial. Listeners completed a 4I2AFC discrimination task, which optimizes performance. On each trial they indicated which of two tone pairs contained a version of the tone that was impoverished (to 22, 20, 18, 16, 14, 12, 10, or 8 bits) relative to the 24-bit standard that filled remaining intervals. Trials were blocked by the tone/source (8 trials/bit-depth). Anti-aliased sawtooth waves occurred at any of several fundamental frequencies (110, 440, 880, and 1,320 Hz) to evaluate whether harmonic spread influenced discrimination. Likewise, sampled instrument tones from clarinet, violin, and piano (at a 440 Hz fundamental) were included to determine if differences in spectral and/or amplitude envelopes were critical to task performance.

Listeners could not reliably detect comparison tones above 12-bits ($d' < 1$). Thus, although exceeding CD (16-bit) quality provides a lower noise floor and greater dynamic range, at comfortable listening levels there do not appear to be clear perceptual benefits, at least not under our psychoacoustic testing conditions. At the lowest comparison bit-depths, discrimination varied with fundamental frequency, with improved performance as harmonics spread. Natural instrument comparisons were more discernible than their synthesized counterparts at lower bit-depths, particularly when the tones were more dynamic, with significant portions that were low in amplitude (i.e., piano and clarinet). This may be due, in part, to differential quantization error, where equivalent DAC integer changes translate to greater log/dB changes at low amplitudes. Future research should focus on longer stimuli that are dynamically variable and/or generally low in amplitude to build upon these primary findings.

Author e-mail: hallmd@jmu.edu

Back to Poster Session

26 Perceptually-based considerations regarding the digital representation of amplitudes

Michael D. Hall* James Madison University

Technological advances have enabled continued improvements in digital audio resolution through increased sampling frequency and bit-depth, respectively impacting the representation of maximum frequency and amplitude changes. Increasing bit-depth reduces quantization error relative to intended amplitudes and reduces the noise floor, expanding available signal range. Questions have been raised as to whether exceeding 16-bit resolution is useful for listening. After all, the noise floor appears low at 16-bit resolution, and frequencies within hearing limits are captured at 16- and 24-bit standards. Furthermore, compression algorithms (like .mp3) diminish bit-depth for samples where the excluded information was likely previously masked (e.g., Moore, 2013). Recently, our laboratory (Stone, 2023) even demonstrated that, at reasonable volumes, listeners cannot discriminate 16- and 24-bit versions of complex tones, despite a paradigm that maximized sensitivity to acoustic details. In fact, listeners could only distinguish bit-depths below 16 at lower signal amplitudes, where quantization error increases.

The current talk explores potential benefits of revisiting how we represent amplitude in audio files. It will be argued that, currently, 16-bit resolution is sufficient for listening experiments at comfortable levels. Further benefits could be attained if changes in amplitude were coded in perceptually relevant terms. It will be demonstrated (by resynthesis after rounding every sample to the nearest portion of a dB) that if DAC integers corresponded to incremental changes on a log scale rather than directly to changes in voltage, then < 15 bits would likely be sufficient to represent complex signals without relevant information loss and without meaningful gain in the noise floor. Such a method also should reduce detectability of quantization error at lower amplitudes. Implications will be discussed with respect to related future investigation, as well as reduced file sizes and maximal storage/portability, including benefits for internet-based experiments as researchers seek more diverse and representative samples of listeners.

Author e-mail: hallmd@jmu.edu

Back to Poster Session

27 Strategies for fitting response time data in psychoacoustic experiments: Insights from a noisy exemplar approach

Nathan F. Gillespie* University at Albany, SUNY
Gregory E. Cox University at Albany, SUNY

Computational cognitive models are powerful tools that explain how representations of perception and memory are transformed into observed outcomes—such as choice and response time—at the level of individual participants. Such models have been used across domains in cognitive science to build integrative theories of phenomena like categorization, recognition, and attentional learning. Many of these models, however, were developed to account for data from visual tasks, so it remains unclear whether the same mechanisms apply in the auditory domain. Thus, theories of auditory memory have not benefited from modeling techniques to the same extent as have theories of visual memory. To address this gap, we fit data from three experiments on timbre perception and memory using the Exemplar-Based Random Walk model (EBRW; Nosofsky & Palmeri, 1997), a prominent model of visual recognition and classification. Using multidimensional scaling, we inferred the dimensions by which each participant represented timbres and embedded these representations into the EBRW model to predict accuracy and response time on a subsequent timbre recognition memory task. The EBRW model jointly accounted for the effects of perceptual similarity, serial order, and list homogeneity—revealing that the same psychological representations underlie both timbre perception and memory. We discuss strategies for addressing the unique challenges of adapting mathematical models of visual cognition to the auditory domain and highlight the utility of applying such models in revealing connections between seemingly disparate aspects of auditory cognition.

Author e-mail: ngillespie@albany.edu

Back to Poster Session

28 Culture influences multisensory emotion perception in bicultural bilinguals

Ashley Chung-Fat-Kim — Northwestern University
Peiyao Chen — Haverford College
Viorica Marian* — Northwestern University

Emotions play a vital role in our lives and are expressed through facial and vocal expressions. Previous research has shown that individuals from Western cultures focus more on the visual modality (i.e., facial expressions), whereas individuals from Eastern cultures focus more on the auditory modality (i.e., tone of voice) when evaluating the emotions of others. Despite the number of bicultural bilinguals in the world, it remains unknown whether individuals who identify with two or more cultures consisting of different social norms process multisensory emotions from each culture differently. In two experiments, we examined multisensory emotion perception in 41 Chinese-English bicultural bilinguals who moved from China to the U.S. (31 females, MAge = 24.07 years, SDAge = 3.53) and in 31 bicultural bilinguals who live in a multi-ethnic country (i.e., Singapore; 21 females, MAge = 24.00 years, SDAge = 3.33) using an emotion recognition task. In this task, participants were presented with audiovisual stimuli of facial expressions (Asian or Caucasian faces) and emotional speech (Mandarin or English pseudo-sentences) and asked to judge the facial or vocal emotion while ignoring the other modality. Across modalities, the emotions were either the same (e.g., happy face paired with a happy voice) or different (e.g., happy face paired with a sad voice). The same voice was always paired with the same face to maintain consistency in identity. Additionally, the auditory and visual stimuli were from the same culture (e.g., Asian face with Mandarin speech and Caucasian face with English speech). In both experiments, we found that bicultural bilinguals focused more on facial expressions when evaluating Western audiovisual emotional input and focused equally on vocal and facial expressions when evaluating Eastern audiovisual emotional input. We conclude that culture plays an important role in shaping the way bicultural bilinguals perceive and judge the emotions of others.

Author e-mail: v-marian@northwestern.edu

Back to Poster Session

29 The role of basal ganglia in Chinese garden-path sentence comprehension processing

Yingying Tan*	Shanghai International Studies University
Ran Zhang	Shanghai International Studies University
Xiaolin Zhou	Shanghai International Studies University
Zheng Ye	Chinese Academy of Sciences

Incremental language comprehension often involves resolving temporary ambiguities and extracting the correct meaning. While previous studies have highlighted the role of cortical regions like the left IFG, MTG, and MFG in ambiguity resolution, the role of subcortical structures, such as the basal ganglia (BG), remains less understood. This study explored the involvement of the BG in processing ambiguous Chinese sentences.
Fifty healthy participants underwent fMRI while reading Chinese sentences with ambiguous verb phrase (VP) + noun phrase (NP) structures. Ambiguous sentences, like "bite hunter de dog" (interpretable as either “the dog that bit the hunter” or “someone bit the hunter’s dog”), were compared to matched unambiguous sentences. Ambiguity (ambiguous vs. unambiguous) and structure type (NP-preferred vs. VP-preferred) were manipulated, resulting in four conditions. Behavioral tests assessed participants' WM and inhibition abilities.

Results confirmed that ambiguity resolution activates the frontal-temporal language network. Participants were less accurate following NP-preferred structures, which activated the bilateral supramarginal gyrus and left precuneus, suggesting involvement in processing canonical word order. A significant interaction showed higher activation in the bilateral MFG, right superior frontal gyrus, and left precentral gyrus for NP-preferred ambiguous sentences, indicating syntactic reanalysis when canonical order is violated. Crucially, ROI analysis revealed that the left caudate, particularly its anterior part, was sensitive to ambiguity resolution. These findings suggest that both the cortex and BG are involved in ambiguity resolution during sentence processing.

Author e-mail: tyy3737@gmail.com

Back to Poster Session

30 Shocking results: Impacts of auditory-tactile training on auditory perceptual learning

Chelsea Joyner* Emporia State University, Kansas State University
Matthew G. Wisniewski Kansas State University

Perceptual experiences tend to be multisensory ones, yet multisensory perceptual learning is not as often studied as learning in the individual senses (i.e. vision, audition, or tactile). The current experiments looked at auditory-tactile training in comparison to unimodal auditory-only training on auditory perceptual tasks. A series of experiments were completed in which participants were trained with unimodal auditory-only, or synchronous multimodal auditory-tactile sounds in temporal discrimination and categorization tasks. Some experiments tested generalization to auditory-only detection tasks or used auditory-only testing of the same task used during multisensory training. Sensitivity (d') and/or thresholds were measured and compared across training conditions, and over time. Experiments found an increase performance over time, demonstrating learning. However, there was only one occasion in which auditory-tactile learning occurred at an accelerated rate over time compared to the auditory-only training. We discuss the possible mechanisms of multisensory learning benefits, and why they may not have been found in these studies. In the applied realm, this research may inform optimizing perceptual training methods that are relevant to real-world perception issues (e.g., reactions to familiar sounded alarms; rehabilitation in auditory processing disorders).

Author e-mail: cjoyner1@emporia.edu

Back to Poster Session

31 Emotions and altered consciousness in voice-induced synesthesia

Cathy Lebeau*	Université du Québec à Montréal
François Richer	Université du Québec à Montréal

In voice-induced synesthesia, voices trigger atypical supplementary perceptions (colors, shapes, touch) in relation to their acoustic features. We present three cases suggesting that the limbic system plays important roles in voice-induced synesthesia. The three cases reported that all voices induce visual or somatosensory synesthesia. Also, their synesthesia change when voices elicit emotions, and they reported other synesthesia triggered by emotions and visceral sensations. Furthermore, these cases reported consciousness alterations in everyday life, including dissociations (time, space and body alterations, out-of-body experiences, Capgras syndrome) and other unusual beliefs or hallucinations (mind reading, premonitions, impression of presence) but no psychiatric disorder. All these phenomena have been linked to limbic system or insular cortex alterations. These cases suggest a limbic/insular hyperexcitability in voice-induced synesthesia.

Author e-mail: cat_lebeau@hotmail.com

Back to Poster Session

32 Congruency advantage in multisensory processing is task-specific

Kayla Phan*	University of San Diego
Ellis Noteboom	University of San Diego
Susan Geffen	Occidental College
Chris Robinson	The Ohio State University, Newark
Laura Getz	University of San Diego

Creating a unified perception of the environment requires rapid organization of complex inputs across sensory modalities. Research using semantically congruent/incongruent stimuli pairs have shown that multimodal presentation influences task performance; congruency across modalities enhances performance, whereas incongruency interferes (Beck et al., 2023). To examine when congruency affects audiovisual processing, this study used a task with audiovisual animal/vehicle pairs of five congruency levels: "semantic response congruent" (SRC) presented stimuli representing the same identity (dog image, bark sound), "response congruent" (RC) presented the same category but different identities (dog image, meow sound), "incongruent" presented different categories (dog image, train sound), "baseline" presented irrelevant secondary stimuli (dog image, pinball sound), and "unimodal" presented one modality (dog image, no sound).

Across four experiments, participants completed two blocks, one responding to visual stimuli and one to auditory. Experiment 1 (N=21) used sequentially presented stimuli with one modality preceding by 300ms and varying onsets between trials. Experiment 2 (N=24) used simultaneous presentation with varying trial onsets. Experiment 3 (N=19) had simultaneous presentation with a fixed 500ms between trials. Experiment 4 (N=52) removed the fixation and response screens, directly replicating Beck et al. (2023).

For Experiments 1-3, effects of congruency and incongruency appeared on the visual response block. Any auditory stimulus sped up responses relative to unimodal trials but at the cost of accuracy on incongruent and baseline trials. There were few differences on the auditory response block until Experiment 4, where auditory accuracy decreased on incongruent trials (replicating previous results) but also SRC and unimodal conditions. Additionally, response speed was lower on SRC and higher on incongruent relative to baseline (replicating previous results) but was also higher on RC trials. Inconsistent effects suggest the need for future EEG/ERP analyses to illuminate the time course of this audiovisual cognitive processing.

Author e-mail: kaylaphan@sandiego.edu

Back to Poster Session