APCAM 2025

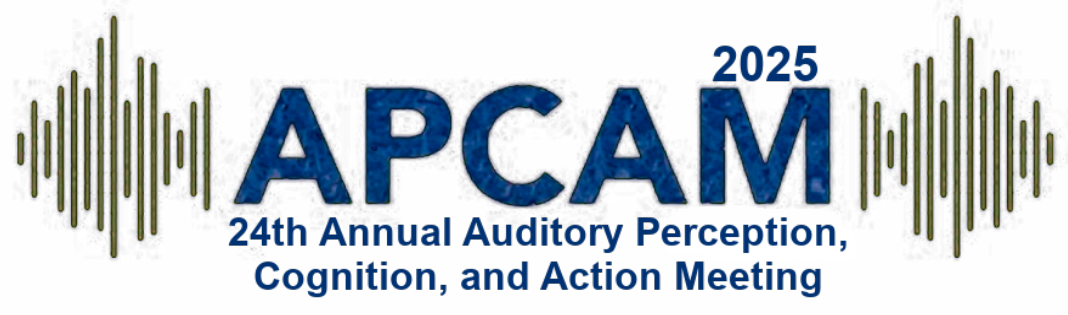
24th Annual Auditory Perception, Cognition, & Action Meeting

Thursday, November 20th, 2025

Sheraton Denver Downtown

Denver, CO

8:00am - 5:00pm



Welcome to the 24th annual Auditory Perception, Cognition, and Action Meeting (APCAM 2025)! 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 and diverse perspectives, 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 2025 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 auditory science. 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 66th Annual Meeting of the Psychonomic Society, APCAM is indebted to the Psychonomic Society for material support, and we also 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 Daniel Tollin, 18 spoken sessions, and 22 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 2025 Organizing Committee

Timothy L. Hubbard (Chair)

J. Devin McAuley

Kathleen C. McCulloch

Kristopher J. Patten

Peter Q. Pfordresher

Sponsors and Donors

Thanks to our sponsors and donors!

Gold Sponsors

PSYCHONOMIC SOCIETY®

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

FINDING FIVE

https://www.findingfive.com/

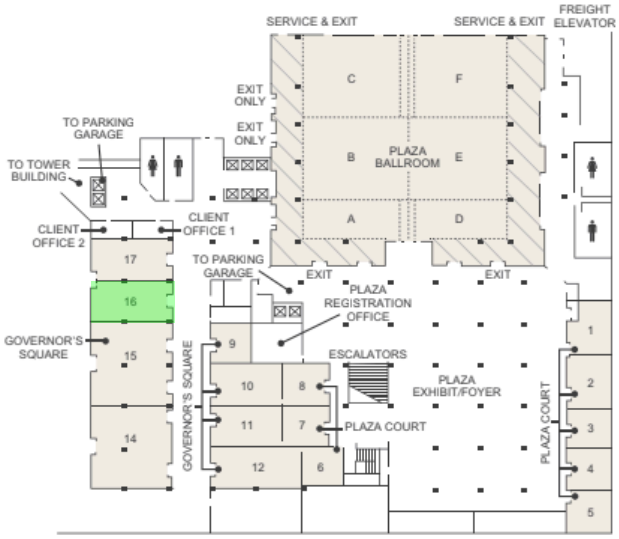
Special thanks to Finding Five for sponsoring the DEI Travel Scholarship!

Member donations were made by:

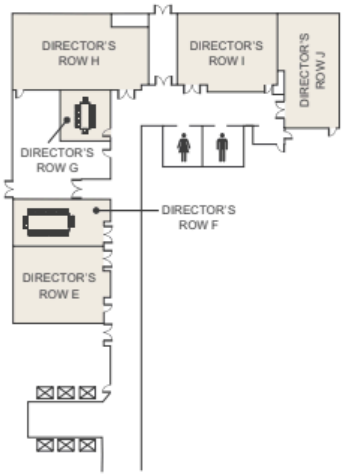
- Leah Fostick
- Sara Haigh
- Tim Hubbard
- Erica Knowles
- Mel Mallard
- Devin McAuley
- Jenny Roche
- Joel Snyder
- Katie Tetzloff

Map

Spoken Sessions - Plaza Building, Governor's Square 16
Concourse Level

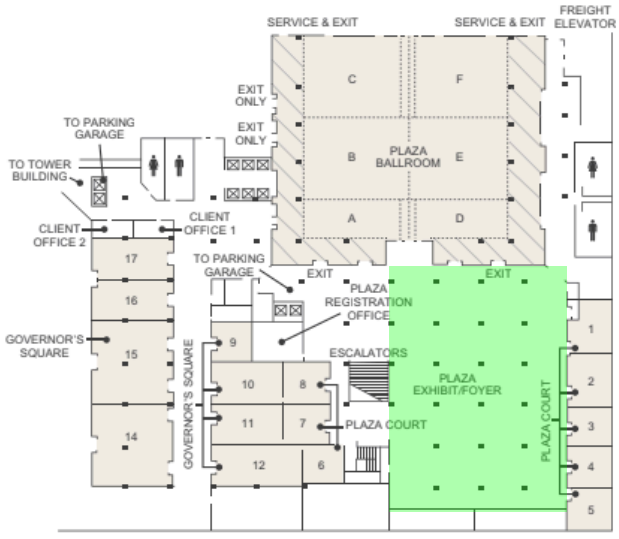


Concourse Level

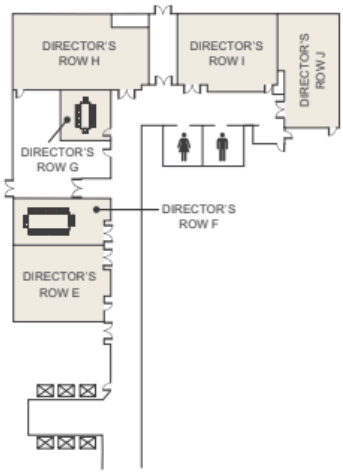


Lobby Level

Poster Session - Plaza Building, Exhibit/Foyer
Concourse Level



Concourse Level



Lobby Level

Full Schedule

8:00	**Registration**	
8:10	**Welcoming Remarks**	
8:15	**Music Cognition & Cross-Modal Perception**	
8:15	The development of musicality: Comparing developmental trajectories of perception and self-report	**Dylan Christie**, U. of Nevada, Las Vegas
8:30	Absolute pitch in involuntary musical imagery	**Matthew G. Evans**, UC Santa Cruz
8:45	Exploration of music as a tool for memory encoding and retrieval	**Isabella L. Ramirez**, U. of San Diego
9:00	Audiovisual integration in virtual reality: Physical vs virtual sound sources	**Maggie McCracken**, U. of Utah
9:15	The re-wiring brain: Cross-modal neuroplasticity in hearing loss	**Anu Sharma**, U. of Colorado, Boulder
9:30	**Break**	
9:45	**Timing, Cognitive Control, & Auditory Scene Analysis**	
9:45	Challenging listening conditions increase the persistence of cognitive control in an auditory flanker task	**Lauren Petley,** Clarkson U.
10:00	The contribution of temporal fine structure to individual differences in auditory temporal order judgment	**Leah Fostick**, Ariel U.
10:15	Neural markers of implicit metric processing in 6-to-10-year-old children	**Katerina Drakoulaki**, Mount Holyoke College
10:30	An open source set of naturalistic, binaural, non-verbal audio stimuli from North American home environments	**Aiden C. Iveris**, UC San Diego
10:45	**Break & Poster Set-Up**	
11:00	**Poster Session**	

12:15	**Lunch**	
1:30	**Keynote**	
	Impact of extended high-frequency hearing loss on neural and behavioral temporal processing and binaural hearing	**Daniel J. Tollin** U. of Colorado School of Medicine
2:00	**Business Meeting**	
2:20	**Break**	
2:30	**Auditory Memory & Imagery**	
2:30	Computational cognitive modeling of individual differences in auditory recognition	**Nathan F. Gillespie**, U. at Albany, SUNY
2:45	Neural indicators of misophonia related cognitive load	**Michael A. Tollefsrud**, Kansas State U.
3:00	Humans display delayed generalization of learning over a time scale of minutes on an auditory perceptual task	**Beverly A. Wright**, Northwestern U.
3:15	Redefining the distinction between the mind's ear and the inner voice	**Timothy L. Hubbard**, Arizona State U.
3:30	**Break**	
3:45	**Speech & Language**	
3:45	Effects of classroom noise on intelligibility of young children's speech	**Katerina A. Tetzloff**, Stony Brook U.
4:00	Categorical perception of speech or resurrecting a dead horse	**Michael D. Hall**, James Madison U.
4:15	Making the case for aperiodic neural activity as a unique index of listening effort	**Sarah J. Woods**, U. of Utah
4:30	Reducing listening effort with accent exposure	**Kristin Van Engen**, Washington U. in St. Louis
4:45	Prosody, knowledge, and misconceptions: Vocal cues in real-time decision making	**Shruti Kate**, Kent State U.
5:00	**Concluding Comments**	

Poster Session

1-3	Auditory Scene Analysis	
1	Auditory classification decisions: Modeling evidence accumulation for signals in naturalistic environments	**Pierce C. Johnson**, U. at Albany, SUNY
2	Impact of repeated presentations on occlusion detection: Does more experience result in greater perceptual accuracy?	**Mike Russell**, Bellevue U.
3	Auditory-induced vection and motion sickness	**Carson Smith**, U. of Nevada, Reno
4-7	**Clinical/Hearing Loss**	
4	Auditory sensitivity in migraine – What is causing it?	**Sarah M. Haigh**, U. of Nevada, Reno
5	The effect of underfitted hearing aids on neurophysiologic and behavioral outcomes	**Kayla Cormier**, U. of Colorado, Boulder
6	Detection of sounds and auditory discomfort	**Sidney K. Hulsey**, U. of Nevada, Reno
7	Can practice listening with a simulated hearing protection device benefit worker performance while wearing hearing protection?	**Patricia McClain**, Kansas State U. Children's Mercy Hospital
8-10	**Crossmodal Perception**	
8	"Synekinian pairs": Manual-vocal gesture integration in experimental contexts	**Grégory Beller**, HfMT Hamburg
9	Multisensory integration in speech perception	**Paméla Trudeau-Fisette**, U. de Montréal
10	Does endogenous visual cueing influence audiovisual integration?	**Madison Prudencio**, U. New Brunswick, Saint John

11-14	Music Cognition	
11	Quantitatively exploring the relationship between self-reported auditory affective experiences, real-time emotional reactions, and music perception	**David J. Vollweiler**, U. Nevada, Las Vegas
12	Starting on the right note: Exploring a working memory-centered ear-training pedagogy	**Erica R. Knowles**, Berklee College of Music
13	New music consumption habits reshape autobiographical memory to reminiscence tiers	**Justine Canio**, U. of San Diego
14	Evaluating potential influences of background music on working memory capacity	**Carter Peery**, James Madison U.

15-21	Speech & Language	
15	Transfer of statistical learning from speech perception to production generalizes to reading	**Kyle Huffaker**, Indiana U., Bloomington
16	CNV as a neural signature of speech production modulated by complexity and planning	**Mohsen Parsa**, Arizona State U.
17	Cognitive control during speech recognition in noise in ADHD	**Victoria Holcomb**, Auburn U.
18	Generalization of statistical learning of vowels in speech perception and production	**Lin Zhou**, U. of Texas at Austin
19	Exploring decision-making processes in speech production	**Jihee Renee Kang**, Arizona State U.
20	Direct brain recordings show strengthening prioritization of speech over music	**Rajvi Agravat**, U. of Texas at Austin
21	Speech envelope modulation detection across frequency bands	**Milad Yousefi**, U. of Illinois, Urbana-Champaign

2025 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.

- Milad Yousefi (University of Illinois Urbana-Champaign)

Talk Abstracts

Music Cognition & Cross-Modal Perception (8:15-9:30)

8:15 The development of musicality: Comparing developmental trajectories of perception and self-report

Dylan Christie*	University of Nevada, Las Vegas
David J. Vollweiler	University of Nevada, Las Vegas
Rose De Kock	University of Nevada, Las Vegas
Joel S. Snyder	University of Nevada, Las Vegas

Musicality is comprised of a range of perceptual abilities and motivational factors. While some aspects of musicality may depend on listening experiences during childhood, others may remain stable throughout development. Here we ask whether the development of behavioral performance in musical tasks are similar to each other and to the development of self-reported musicality. Children (N = 241) and adults (N = 246) completed tasks measuring rhythm perception (internal beat perception, beat sensitivity, meter sensitivity), pitch perception (sensitivity to key and harmony, major/minor chord identification), emotion identification in music, and sensitivity to speech prosody. Children also completed the child Musicality Index (cMI) to assess self-reported musical interest and motivation. Performance on all tasks improved with age, however we observed a plateau during adolescence for some tasks (meter) and continued improvement into adulthood for others (key and harmony). By contrast, self-reported musicality (cMI scores) remained stable throughout childhood. When controlling for age, performance was correlated across all perceptual tasks with the exception of meter sensitivity, which only correlated with sensitivity to beat and prosody. Musicality correlated modestly with both beat perception tasks and musical emotion identification after controlling for age. Overall, these findings suggest that while interest in and motivation to make music is developmentally stable and does explain some individual variation in music perception skills, all children steadily improve their rhythm and pitch abilities with age, albeit with subtle differences in trajectories for different skills. These results highlight the value of combining behavioral and self-report measures across the same individuals to gain a more complete picture of musical development, and they underscore that musicality cannot be reduced to a single trait.

Author e-mail: chrisd6@unlv.nevada.edu

Back to Schedule

8:30 Absolute pitch in involuntary musical imagery

Matthew G. Evans*	University of California, Santa Cruz
Pablo Gaeta	University of California, Santa Cruz
Nicolas Davidenko	University of California, Santa Cruz

Memory for isolated absolute pitches is extremely rare in Western, English-speaking populations. However, past research has found that people can voluntarily reproduce well-known songs in the original key much more often than chance. It is unknown whether this requires deliberate effort or if it manifests in involuntary musical imagery (INMI). Participants (N = 30, convenience sample) were surveyed at random times over a week and asked to produce a sung recording of any music they were experiencing in their heads. We measured the "pitch error" of each recording to the nearest semitone by comparing participants' recordings to the original song. We found that 44.7% of recordings had a pitch error of 0 semitones, and 68.9% of recordings were within ± 1 semitone of the original song. Our results provide novel evidence that a large proportion of the population has access to accurate absolute pitch memory, as revealed in their INMI.

Author e-mail: magievan@ucsc.edu

Back to Schedule

8:45 Exploration of music as a tool for memory encoding and retrieval

Isabella L. Ramirez*	University of San Diego
Laura Getz	University of San Diego

Melodies as mnemonic devices are a common strategy to support memory, particularly in educational settings. Regardless, findings are actually mixed regarding whether music enhances memory in the short term. In Experiments 1 and 2, we evaluated how well-known and unfamiliar melodies influenced the memorization of unrelated word lists compared to a spoken-word control. After each list was presented, participants completed a free recall and old/new recognition task. Overall, spoken presentations outperformed the sung conditions.

Experiment 3 shifted focus to repetition, using a single melody presented multiple times to examine whether increased exposure would affect memory outcomes. Results showed that repeated melodies produced accuracy levels similar to spoken words.

Experiments 4 and 5 aimed to further our understanding of how melody complexity and familiarity impact memory encoding. We created unfamiliar melodies matched in complexity to the well-known melodies used in Experiments 1-3. Results of Experiment 4 showed that performance on these melodies was similar to the complex familiar melodies, highlighting the importance of complexity over familiarity in memory performance. In Experiment 5, participants were presented with word lists sung to unfamiliar complex melodies. Afterward, they were instructed to listen to one of these melodies without lyrics several times per day for a week. When they returned to the lab, they completed the same memory task using that now-familiar melody amidst unfamiliar melodies. This within-subject design allowed us to compare memory performance for the same melody when it was unfamiliar versus when it became familiar, showing improved performance in musicians but not non-musicians.

Together, these five experiments suggest that while musical mnemonics are widely used, familiar melodies may initially interfere with memory performance, possibly due to competition between new material and previously learned lyrics. In the context of short-term memory, music may not always serve as an effective aid.

Author e-mail: izzie.ramirezz@gmail.com

Back to Schedule

9:00 Audiovisual integration in virtual reality: Physical vs virtual sound sources

Maggie McCrackeen*	University of Utah
Adhithya Narayanan	University of Utah
Sarah Creem-Regerh	University of Utah
Jeanine Stefanucci	University of Utah

Spatial audio is increasingly used in virtual reality (VR), yet differences in the localization of virtual sounds may alter how auditory and visual cues are integrated. Unlike physical sound sources, virtually-rendered audio may be affected by standardized head-related transfer functions (HRTFs), hardware limitations, and head-tracking latency, potentially introducing perceptual discrepancies. In this study, we examined how physical versus virtual sounds influence audiovisual distance perception and integration in VR. Participants judged the location of audiovisual targets with sounds produced either by physical speakers or by virtual rendering using the HTC Vive Pro. Psychometric modeling revealed that virtual sounds yielded less precise and less accurate audiovisual integration compared to physical sounds, with effects moderated by stimulus distance. These findings highlight the importance of considering auditory source fidelity in VR audio design and the need for future research on cross-modal integration in virtual environments.

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

Back to Schedule

9:15 The re-wiring brain: Cross-modal neuroplasticity in hearing loss

Anu Sharma* University of Colorado, Boulder

A basic tenet of neuroplasticity is that the brain will re-organize following sensory deprivation. Compensation for the deleterious effects of hearing loss include alterations in cortical functional dynamics such as engagement of additional or alternative cortical networks for listening to degraded auditory input. Cross-modal plasticity is a form of cortical re-organization associated with hearing loss. Cross-modal plasticity occurs when an intact sensory modality recruits and repurposes cortical resources from a deprived sensory modality. Our experiments using high- density EEG in human subjects show evidence of recruitment of higher-order auditory cortical areas by visual and somatosensory modalities in both developmental deafness and age-related hearing loss. We see evidence of cross-modal reorganization from vision and somatosensation in all degrees of hearing loss (i.e., from mild-moderate hearing loss to profound deafness) and in both congenital and acquired hearing loss. Across all subject groups, cross-modal plasticity is inversely related to scores on clinical tests of speech in noise perception suggesting that recruitment by other sensory modalities may influence the variability in outcomes in persons with hearing loss. Cortical re-organization in age-related hearing loss is also associated with effortful listening, decreased cognitive reserve and neurocognitive decline. Importantly, if cross-modal plasticity reversed with appropriately fitted hearing aids or cochlear implants, it results in good clinical outcomes coinciding with gains in cognitive performance and speech perception. Overall we find that cross-modal plasticity represents a neuronal process that is dynamic and reversible, resulting from existing sub-threshold multisensory inputs to sensory cortices and top-down influences on inhibition, which adaptively interacts with restoration of sensory loss. The versatile nature of cross-modal plasticity can be exploited clinically for improving the outcomes after neurosensory restoration.

Author e-mail: anu.sharma@colorado.edu

Back to Schedule

Timing, Cognitive Control, & Auditory Scene Analysis (9:45-10:45)

9:45 Challenging listening conditions increase the persistence of cognitive control in an auditory flanker task

Lauren Petley* Clarkson University

Flanker tasks are popular tools for studying executive function, notably conflict resolution – suppressing an incorrect response in order to produce the correct one. Applying sequence analyses to these tasks additionally reveals the time course of cognitive control, typically revealing two effects. If the next trial also involves conflict, it is resolved more readily due to preparation, producing faster response times. If it doesn't, responses are slowed versus scenarios where no conflict was encountered. In a series of three experiments, we previously validated an auditory flanker task that uses spoken directional commands as stimuli. To explore the potential influence of masking, trials were also included that imposed masking effects. Using three different vocal configurations, 1) a woman versus a boy (F0 difference = 32.1 Hz), 2) a woman versus a man (F0 difference = 93.1 Hz), and 3) a man versus a girl (F0 difference = 122.8 Hz), we demonstrated that conflict was greatest when the spectral separation between voices was small. The present analysis applied a sequence analysis to these data to explore differences in the persistence of cognitive control. With the smallest amount of spectral separation, conflict significantly slowed the response on the next trial, compared to trials that only contained masking. This conflict effect was observed regardless of the category of the next trial. Increasing the spectral separation between the talkers reduced this conflict effect. Specifically, with moderate separation, the effect of conflict was modified by the category of the next trial, making it inconsistent, while only a near-significant conflict effect was observed with the greatest amount of spectral separation. Combined with our previous findings, this analysis reveals that challenging listening conditions not only increase the conflict that is evoked by simultaneous talkers, they lead to more persistent exertion of cognitive control in anticipation of future conflict.

Author e-mail: lpetley@clarkson.edu

Back to Schedule

10:00 The contribution of temporal fine structure to individual differences in auditory temporal order judgment

Leah Fostick* Ariel University

Substantial variability exists in basic auditory skills, even among individuals with clinically normal hearing. This study investigates the factors underlying individual differences in spectral temporal order judgment (TOJ), a task where listeners identify the sequence of two brief tones. We previously identified three distinct performance patterns: High-Level Performance (High-LP), with consistently high accuracy at all inter-stimulus intervals (ISIs); Mid-Level Performance (Mid-LP), where accuracy improves gradually with increasing ISI; and Low-Level Performance (Low-LP), which shows poor performance even at long ISIs. We propose that temporal fine structure (TFS) may be the key acoustic cue that differentiates these strategies. TFS refers to the rapid fluctuations in a sound wave, and its encoding in the auditory system supports pitch perception and sound localization. We conducted three experiments to test this hypothesis. Experiment 1 found that increasing the frequency separation between tones, which reduces TFS cues, led to a decrease in High-LP and an increase in Mid-LP. Experiment 2 showed a similar shift: as the absolute frequency of tones increased (where TFS cues are less robust), the proportion of High-LP decreased, while Mid-LP and Low-LP increased. Finally, Experiment 3 demonstrated that masking TFS cues with a narrow-band noise significantly reduced High-LP and increased Mid-LP. Low-LP was largely unaffected across all experiments. These findings collectively suggest that individual differences in spectral TOJ performance are linked to underlying differences in the perception and use of TFS cues. High-LP may be associated with effective use of TFS, while Mid-LP may reflect a greater reliance on time-based cues. This research highlights the critical role of TFS in auditory temporal processing and provides a sensory-level explanation for the observed variability in perceptual strategies. Future work will explore how cognitive abilities like working memory and attention might influence TFS utilization.

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

Back to Schedule

10:15 Neural markers of implicit metric processing in 6-to-10-year-old children

Katerina Drakoulaki*	Mount Holyoke College
Mara Breen	Mount Holyoke College

Children's rhythmic and metric processing abilities are shown to predict reading skills, but the connection between these processes is unclear. This study explores the electrophysiological development of metric processing in children ages 6-to-10, to understand its relationship to reading skill.

We adapted an imagined meter paradigm validated in adults to determine the maturational trajectories of neural markers of implicit metric coding. If children's metric encoding matures during this age range, we predict that electrophysiological correlates of implicit metric structure – the N1 and late metric negativity (LMN) in response to metrically strong sounds – will increase with age. High-density electroencephalography (EEG) data were recorded from N = 74 children (out of a planned 120), while they listened to up to 40 one-minute-long trials of an isochronous stream of undifferentiated tones as imagined groups of three or four. The 40 trials included 10 for each combination of listening pattern (triple, quadruple) and presentation tempo (fast = 450ms ISI; slow = 625ms ISI).

Across four 15-month age cohorts, we replicate previous findings regarding the development of the morphology of auditory evoked potentials (AEPs) in this age range: Similar to previous work, we find the P1 component present in 6-year-olds, with decreasing latency across age cohorts; moreover; we find N1 peaks for the two oldest age cohorts. In addition, we show the presence and maturation of LMN morphology, with increasing LMN differentiation of metric strength across age. Finally, we report a negative correlation between the LMN effect and reading comprehension scores: children with better reading skill had larger negative responses to strong beats relative to weaker beats. These results demonstrate that acoustic processing of metric structure matures between ages 6-10, and that implicit metric coding emerges during this window. Moreover, the results offer insight into the relationship between meter perception and reading comprehension.

Author e-mail: adrakoulaki@mtholyoke.edu

Back to Schedule

10:30 An open source set of naturalistic, binaural, non-verbal audio stimuli from North American

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

Research testing perception of environmental sounds is often limited by lack of accessible, high-quality recordings of verifiable and naturalistic acoustic events. Researchers often rely on commercially produced, artist-designed audio libraries. In many of these sound libraries some putative sounds are not veridical (e.g., stones on straw for horses galloping; Costanzo, 2020); typically these libraries are inadequately documented for research. Alternatively, some researchers laboriously compile stimuli from idiosyncratic sources with varying recording equipment quality and methods. Here we present an open source stimuli library of 250+ genuine sounds from everyday household events, recorded by five researchers in eight different homes, and supplemented with a rich descriptive metatag database and corresponding photographic records. We additionally share procedures developed to maximize ecological relevance in the resulting stimuli.

Collection methodology prioritized perceptual fidelity from an embodied perspective. Audio was collected with a DPA 4560 Binaural Headset. The paired, omnidirectional microphones have a frequency sensitivity range exceeding that of most humans, and were placed at the entrance of researchers' ear canals, capturing the effects of their HRTF on incoming sounds. Sounds were recorded on a Zoom F3 Field recorder at 32-bit float, with uncompressed dynamic range and accurate relative sound levels within and across recordings. Researchers were trained to capture specific target sounds as part of longer, human-centered action sequences, following their typical routines when engaging with household objects. The resulting audio captures changes in body positioning while interacting with objects (e.g., looking towards an alarm). Written annotations and images accompany each sound.

Time permitting, we will describe a project to norm adults' accuracy in identifying and characterizing various sounds in the dataset. We anticipate this sound library will be of interest to researchers investigating environmental sound perception, particularly when ecological accuracy and knowledge of the real-world attributes of sound sources are desired.

Author e-mail: aiveris@ucsd.edu

Back to Schedule

Keynote (1:30-2:00)

Impact of extended high-frequency hearing loss on neural and behavioral temporal processing and binaural hearing

Daniel J. Tollin* University of Colorado School of Medicine

Successful binaural hearing, including speech understanding in noise, depends on precise neural encoding of ear-specific cues such as interaural level and timing differences. Even with clinically normal audiometric thresholds, reduced access to these cues can produce substantial listening difficulties. Emerging evidence implicates peripheral dysfunction, particularly elevated extended high-frequency (EHF; >8 kHz) thresholds, in degraded suprathreshold performance on temporal and spatial tasks. Here, we examined whether EHF hearing loss is associated with deficits in temporal processing and binaural hearing, and with altered neural biomarkers of sound encoding across the auditory pathway.

35 participants with normal bilateral thresholds from 250–4000 Hz completed behavioral assessments of temporal fine-structure sensitivity and spatial speech-in-noise perception. Electrophysiological measures included the summating-potential to action-potential ratio (SP/AP), auditory brainstem response (ABR) with the binaural interaction component (BIC), speech-evoked frequency-following response (sFFR), and the interaural phase-modulation following response (IPM-FR).

Preliminary analyses indicate that listeners with EHF loss (N = 16) perform significantly worse on spatial speech-in-noise and temporal fine-structure tasks than those with normal EHF hearing (N = 19). Spectral analyses of evoked responses in the EHF loss cohort further reveal degraded subcortical phase locking (reduced sFFR), diminished ABR BIC, and weaker cortical representations of binaural cues (attenuated IPM-FR), despite normal thresholds within the frequency ranges probed by these assays. Additionally, individuals with EHF loss showed elevated SP/AP ratios, consistent with cochlear synaptopathy - the loss of auditory nerve fiber to hair cell synapses.

These converging behavioral and electrophysiological effects suggest that EHF impairment is a sensitive indicator of subclinical listening difficulties and may reflect “hidden” hearing loss arising from cochlear synaptopathy. Incorporating EHF assessment and objective neural markers into clinical and research protocols could improve early detection of binaural and temporal processing deficits that are not apparent on standard audiometry.

Author e-mail: daniel.tollin@cuanschutz.edu

Back to Schedule

Auditory Memory & Imagery (2:30-3:30)

2:30 Computational cognitive modeling of individual differences in auditory recognition

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

We present four experiments that examine perception and memory for a novel set of auditory stimuli, using multidimensional scaling and cognitive modeling to clarify how individual people perceive and recognize these items. The stimuli are auditory "textures" constructed by adjusting the distribution of power across upper frequency bands. In Experiment 1, people rated similarity between pairs of stimuli; in Experiments 2 and 3, they also engaged in a recognition memory task using the same stimuli. In Experiment 4, they did all the same tasks from the first three experiments, and rated stimuli for distinctiveness. Multidimensional scaling suggested the stimuli were perceived along three dimensions, a result which replicated across all four experiments. Similarity ratings, recency, and list homogeneity predicted recognition performance, but distinctiveness ratings did not. Individual differences in these effects were accommodated by the Exemplar-Based Random Walk model (Nosofsky and Palmeri, 1997)—extending prior work (Visscher et al., 2007) to show that memory and attention processes in the auditory domain are fundamentally like those in the visual domain, though particularly strong recency effects in the auditory domain may be due to echoic memory. We conclude by discussing how the stimuli introduced in this talk can be used as "building blocks" to test hypotheses about perception and memory for complex, naturalistic sounds such as speech or music while retaining tight experimental control.

Author e-mail: ngillespie@albany.edu

Back to Schedule

2:45 Neural indicators of misophonia related cognitive load

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

Misophonia is characterized by disproportionate emotional and physiological responses to specific “trigger” sounds, often accompanied by self-reported difficulty maintaining focus during exposure. This research sought to provide objective evidence of such cognitive disruptions by combining behavioral measures of auditory working memory (WM) with electroencephalographic (EEG) indices of neural processing. Fifty-seven participants (27 misophonic, 30 control) completed a pitch-matching WM task under three distractor conditions: silence, neutral sounds, and individualized trigger sounds. EEG was analyzed using a time–frequency approach to assess changes in alpha (8–12 Hz) oscillatory activity during the retention interval. Behaviorally, misophonic participants exhibited reduced overall WM performance compared to controls and a selective decrease in response precision. This suggests degraded memory fidelity under distraction by triggers. From the EEG analysis, misophonic participants showed lower alpha power across conditions, consistent with diminished inhibitory control. These results indicate that misophonia may involve a generalized attentional control vulnerability that amplifies the disruptive impact of emotionally salient auditory stimuli. Implications and directions for future research are discussed.

Author e-mail: matollef@ksu.edu

Back to Schedule

3:00 Humans display delayed generalization of learning over a time scale of minutes on an auditory perceptual task

Ruijing Ning — RIKEN, Japan
Beverly A. Wright* — Northwestern University

Training can induce improvement on both the trained task (learning) and related untrained tasks (generalization). However, the emergence of learning and generalization are not necessarily temporally aligned. For example, generalization can lag behind learning. Delays to generalization over time scales of hours and days have been reported previously. Here we document a delay to generalization over a time scale of minutes in human auditory perceptual learning. Thus, delayed generalization appears to be a feature of multiple stages of memory formation. We also document delayed generalization both with a 15-min break between training offset and testing onset and with that break filled with additional training. This outcome suggests that the timing of the emergence of delayed generalization may be more influenced by training onset than training offset. In addition, this outcome aligns with the idea that there is a refractory period in perceptual learning during which more training within a session beyond some sufficient amount does not affect the magnitude of learning across days. More broadly, the presence of delayed generalization suggests that learning and generalization arise from at least partially separable processes and that a full understanding of generalization requires its assessment at multiple time points after training.

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

Back to Schedule

3:15 Redefining the distinction between the mind's ear and the inner voice

Timothy L. Hubbard* Arizona State University

Auditory imagery has traditionally been described as involving the "mind's ear" (imagery of sounds generated by external objects or events) or the "inner voice" (imagery of sounds generated by one's vocalizations), with the former involving auditory information and the latter involving auditory information and motor (articulatory) information. Findings from studies of the role of subvocalization in judgment of imaged auditory stimuli, reading of written text and musical scores, and imagery in singers are noted, and salient results of studies of neuroimaging of musical materials are highlighted. Early attempts to equate subvocalization with the Inner voice were not successful, as motor activation occurs for imagery of external sounds and for imagery of one's vocalizations. It is argued that the notion of "articulation" should be broadened beyond those processes involved in vocal production to include generation of any type of sound (e.g., fingering a musical instrument, breaking a pane of glass), and this suggests a widening of the definition of "inner voice" to include nonvocal sound production. Thus, imagery of an auditory stimulus involves a covert or internal simulation of an overt or external sound-producing action. Such an approach emphasizes the role of action in sound production and the dynamic nature of imagery, and it is suggested that imagery is more properly conceptualized as a verb rather than as a noun. The redefined distinction between the mind's ear and the inner voice maps onto the distinction between the phonological store and the articulatory rehearsal loop in working memory, is consistent with motor theories of speech perception and music perception, consistent with the possibility of mirror neurons, consistent with the distinction between dorsal and ventral auditory pathways in the cortex, and consistent with an embodied view of cognition in which imaged (and perceptual) content is limited by properties of our sensorimotor systems.

Author e-mail: timothyleehubbard@gmail.com

Back to Schedule

Speech & Language (3:45-5:00)

3:45 Effects of classroom noise on intelligibility of young children's speech

Katerina A. Tetzloff*	Stony Brook University
Sarah E. Yoho	The Ohio State University
Stephanie A. Borrie	Utah State University

Accurate perception of children's speech in classroom environments is essential for effective communication, learning, and social interaction. While children's speech is known to differ acoustically from adult speech, exhibiting higher pitch, less stable articulatory targets, and reduced phonetic contrast, very little work has examined how environmental masking, such as realistic classroom noise, affects its intelligibility. This study investigated how background noise typical of classroom environments impacts adult listeners' perception of children's speech. Thirty-nine young, school-age children between three and eight years old (mean = 5.7 years, 56% female) were recorded producing stimuli that consisted of two-word phrases with low inter-word predictability (e.g., "angry money"). These phrases were presented to unfamiliar adult listeners in two listening conditions: in quiet and in classroom noise at a +3 dB signal-to-noise ratio (SNR). Each child's speech was heard by 10 naïve adult listeners (n = 390), who were instructed to orthographically transcribe what they believed the child said. Intelligibility was measured as percent words correct (PWC) in quiet (PWC_quiet) and in classroom noise (PWC_noise); a penalty score was calculated as the difference between conditions (PWC_quiet – PWC_noise). Linear models were used to evaluate results. As expected, intelligibility of the children's speech was significantly lower in noise compared to quiet ($p < .0001$). On average, intelligibility dropped by 14.6% in noise, but children with higher intelligibility in quiet experienced larger decreases (higher penalty) when noise was added ($p < .01$). This counterintuitive result suggests that intelligibility in quiet does not predict robustness in adverse conditions, as even highly intelligible children may be difficult to understand in noisy classrooms. These findings underscore the importance of optimizing classroom acoustics to ensure that all children can be understood, especially in noisy environments. Future work should explore how specific acoustic properties of children's speech contribute to intelligibility in noise.

Author e-mail: katerina.tetzloff@stonybrook.edu

Back to Schedule

4:00 Categorical perception of speech or resurrecting a dead horse

Michael D. Hall* James Madison University

The utility of categorical perception (CP) for speech has been rightfully questioned (see McMurray, 2022) since it has been shown to be sensitive to slight task changes, as well as for being embroiled in phonetic modularity arguments. Yet, some fundamental concerns about CP remain unaddressed. Related evidence will be reviewed, including of decreased performance under presumably optimal conditions that should highlight acoustic differences (a 4I2AFC task), as well as the fact that time delays have been argued to be relatively unimportant despite not exceeding acoustic trace limits (Schouten & Van Hessen, 2003; Gerrits & Schouten, 2004). A previous publication from our laboratory (Hall & Peck, 2017) also reveals that CP for place of articulation is time-sensitive and context-dependent. Specifically, both categorization and 2AFC discrimination performance shifted with changes in F2 transition direction, and phonemic boundaries mapped to frequencies associated with the subsequent vowel. Furthermore, listeners perceived stimuli more continuously when transitions were lengthened, reporting additional categories.

Then a new study will be summarized that sought to address timing and task concerns. Performance was compared under 2AFC and 4I2AFC conditions for the same listeners, and both tasks were completed for two inter-stimulus intervals (ISI's) to see if timing was critical. A discrimination peak corresponding to the labeling boundary, indicating CP, was obtained in all conditions except for the 4I2AFC task with a 350 msec ISI. Thus, a delay that exceeded echoic memory limits (750 msec) was sufficient to reinstate CP. Furthermore, CP was eliminated in an experiment where listeners responded to corresponding nonspeech bleats, suggesting that CP depended upon phoneme perception. These findings suggest a low-level cognitive basis for the effect that reflects the normal course of perceptual processing. Implications, including the likelihood of CP's continued utility, will be discussed, along with concerns regarding common, but false task assumptions.

Author e-mail: hallmd@jmu.edu

Back to Schedule

4:15 Making the case for aperiodic neural activity as a unique index of listening effort

Sarah J. Woods*	University of Utah
Jack W. Silcox	University of Utah
Brennan R. Payne	University of Utah

Aperiodic (non-oscillatory) neural activity is an important component of the electroencephalography (EEG) signal, reflecting dynamic neurophysiological states that are sensitive to cognitive demands. However, its role in speech comprehension remains largely unexplored. We present EEG data from older adults (N = 48) with varying hearing acuity as they listened to sentences in quiet and in background noise. Aperiodic activity was quantified using spectral parameterization (slope, offset). For comparison, we also present data from younger adults with normal hearing who listened to the same sentences (N = 35). Consistent with increased listening effort, aperiodic activity was sensitive to acoustic challenge: spectral slopes flattened, and offsets decreased when speech was presented in noise relative to quiet. Importantly, for older adults these noise-induced changes in aperiodic activity interacted with hearing acuity, suggesting that these measures index not only environmental difficulty but also individual listener factors. Comparisons across age groups replicate known age-related flattening of the aperiodic slope and reduction in offset, while confirming that both younger and older adults show sensitivity to noise.
The present findings contribute to the growing evidence that aperiodic neural activity is not only a source of meaningful variability that must be accounted for in studies of neural oscillations, but also a functionally important measure in its own right. Aperiodic dynamics are sensitive to effort during speech comprehension, indexing differences related to acoustic challenge, aging, and hearing loss. Interestingly, in our samples, aperiodic measures capture effort-related effects in situations where traditional oscillatory measures, such as alpha power, fail to reveal significant differences. Together, these results establish aperiodic activity as a unique and robust neural correlate of listening effort, with implications not only for language and speech research but also for other cognitively demanding tasks where effort is a critical factor.

Author e-mail: sarah.woods@psych.utah.edu

Back to Schedule

4:30 Reducing listening effort with accent exposure

Kristin Van Engen*	Washington University in St. Louis
Kaitlyn Matthews	Washington University in St. Louis
Drew McLaughlin	Villanova University

Mismatched accents among interlocutors can be a significant source of difficulty in speech communication, particularly when the intended message of a speaker is not correctly understood by a listener. However, even when a speaker is highly intelligible to a listener, unfamiliar accents generally impose greater cognitive demands than familiar accents. Previous work on accent adaptation has shown that listeners can make significant and often rapid gains in measures of speech intelligibility and response time. This presentation will review work from our lab showing that listening effort, as measured through pupillary responses and response times in dual-task paradigms, can also be quickly mitigated through accent exposure. Using single- and multi-session laboratory experiments that have manipulated variables such as feedback, talker variability, and talker intelligibility, this work continues to highlight the challenges, complexity, and promise of adapting to speakers who sound different from us.

Author e-mail: kvanengen@wustl.edu

Back to Schedule

4:45 Prosody, knowledge, and misconceptions: Vocal cues in real-time decision making

Shruti Kate* Kent State University
Jennifer Roche Kent State University

When we listen to someone speak, we attend not only to their words but also to contextual cues that shape interpretation. Vocal prosody—intonation patterns that rise, decline, or fall—signals certainty, doubt, or authority. When combined with provisional language such as hedges (e.g., might) or boosts (e.g., definitely), these cues jointly guide how listeners evaluate meaning. This becomes especially important when statements convey either factual information or common misconceptions (e.g., bats are blind). Misconceptions persist because they sound reasonable, align with prior learning, or resemble scientific authority. Yet little research has examined how prosody and provisional language interact to shape decision making in these contexts. This study evaluated how listeners interpret facts and myths framed with hedges and boosts, presented with prosodic contours that either matched or mismatched the lexical signal (e.g., rising intonation with hedged uncertainty, declining intonation with boosted certainty). Using a mouse-tracking paradigm, we assessed both response accuracy and the real-time dynamics of decision making. Results showed that prosodic and lexical cues jointly bias knowledge judgments. Hedged facts elicited the highest accuracy, suggesting that provisional language can promote critical evaluation when paired with true information. In contrast, hedged myths were most often misclassified, indicating that hedging can reinforce misconceptions when listeners hesitate to reject misinformation. Boosting improved accuracy for myths, suggesting that strong certainty cues may help listeners dismiss falsehoods. Intonation also shaped processing: facts spoken with rising intonation slowed responses, while hedged statements with declining intonation produced the greatest cognitive competition, reflected in larger mouse trajectory deviations. These findings demonstrate that how something is said can be just as important as what is said. Prosody and provisional language act as auditory signals that bias how truth and misinformation are judged in real time.

Author e-mail: skate@kent.edu

Back to Schedule

Poster Abstracts

Auditory Scene Analysis (1-3)

1 That's about the shape of it: A scaling study of similarity in a generated set of melodies

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

While pitch and rhythmic content are crucial for melody perception, whether these features are integral or separable, and their relative importance, remains unclear. We investigated these questions using artificial, 10-note melodies generated by sampling pitch intervals from nine unique intervals (the first and last notes in each melody were identical) and rhythmic durations from a set of ten unique intervals. This ensured identical total duration, note count, and available pitch/rhythmic content across melodies, distinguishing them solely by temporal arrangement. We collected subjective similarity ratings from 260 participants who rated all pairwise combinations of 17 melodies on an unmarked slider anchored by "Not at all similar" and "Completely similar." Objective similarity measures were derived by converting melodies into vectors: raw pitch vectors capturing each note's deviation from starting pitch, raw rhythmic vectors recording note durations, and contour vectors calculated by taking differences between a note and the successive note's pitch and duration. Cosine similarity between melody vectors provided objective measures of raw and contour-based pitch and rhythmic similarity. Multidimensional scaling revealed a three-dimensional perceptual space with melodies distributed spherically without clustering, suggesting uniform similarity throughout the set. Regression analyses showed both pitch and rhythm similarity predicted perceived similarity, but critically, only contour-based measures—not raw pitches or rhythms—were significant predictors. Additionally, perceived similarity correlated with preference judgments—similar melodies tended to be liked or disliked equivalently. Pitch contour and rhythmic contour, defined by successive intervals, drove perception. Crucially, no interactions between pitch and rhythm similarity emerged, indicating pitch and rhythmic information is separable rather than integral. These findings demonstrate melodic perception relies on relative patterns of change in pitch and rhythm operating independently.

Author e-mail: johnsonpierce.c@gmail.com

Back to Poster Session

2 Impact of repeated presentations on occlusion detection: Does more experience result in greater perceptual accuracy?

Mike Russell* Bellevue University

Organisms commonly perceive and act in a cluttered world, and clutter can impair our ability to comprehend the world about us and alter how we move through space. As will be discussed, occlusion detection is likely to be relative and not absolute. To be absolute, occlusion detection must be based solely on the information available on a single trial and independent of prior experience. This is unlikely for various reasons. Instead, it is believed occlusion detection is largely dependent on the observer's experience. As can be imagined, observers likely rely on the change in sound across trials to determine whether a sound source is occluded. Support for the argument that occlusion detection is relative is based on research which revealed judgments of sound source location and movement are influenced by an individual's experience (e.g., training, number of trials, feedback). While additional research revealed individuals can use sound to detect the presence of an obstruct, it appears individuals are limited in their abilities. The present study sought to determine whether the accuracy of occlusion detection is influenced by experience. In the present study, participants listened to 5 audio recordings, each differing in the level of occlusion (0-100%, 25% increments). After listening to each recording, participants indicated whether the sound source was occluded. Participants judged each recording either once, three times, or five times. Assuming experience influences occlusion detection, an increase in experience should result in an increase in perceptual accuracy. It was further expected an increase in experience would result in an increased ability to detect partially occluded sound sources. Discussion will be given to how experience, in general, relates to J.J. Gibson's notions of an ambient energy array and information detection, two of the foundations of his ecological approach to perception and action.

Author e-mail: mirussell@bellevue.edu

Back to Poster Session

3 Auditory-induced vection and motion sickness

Carson Smith*	University of Nevada, Reno
Kate Pellegrino	University of Nevada, Reno
Ruby Snyder	University of Nevada, Reno
Sarah M. Haigh	University of Nevada, Reno

Vection refers to the illusory sensation of self-motion that occurs in the absence of actual self-motion. Vection is frequently reported during everyday experiences such vehicle travel and is increasingly relevant to virtual environments and simulations. Visual stimulation is the most well-established inducer of vection and the same cues also induce motion sickness, suggesting similar underlying mechanisms. However, other modalities such as audition are also capable of producing these percepts but are not well explored in isolation. In the present study, we investigated (1) whether auditory cues, in the absence of visual information, can elicit vection and motion sickness, and (2) what conditions could increase the intensity of vection. Participants (N=40) listened to custom Shepard-Risset glissando tones, with the rate of descending sweeps manipulated to assess if presentation speed influenced the percepts. Some trials also included dorsal neck muscle vibrations, which added noise to the muscle spindle afferents about head orientation and motion. Vection and motion sickness were assessed subjectively using a 10-point Likert scale. Our results showed, first, in the absence of visual stimulation, auditory cues can induce vection relative to baseline (where no auditory cues were present; $p<.01$). Second, when assessing the factors that affected vection, there was no significant effect of speed ($p=.49$) or interaction between speed and vibration ($p=.75$) on vection intensity. Interestingly, there was a main effect of vibration stimulation ($p<.01$) on intensity where trials containing both auditory stimuli, regardless of speed, and neck vibrations induced greater vection than trials without vibrations. Additionally, 17 participants experienced motion sickness during auditory stimulation ($p<.01$). These findings suggest that, without vision, auditory cues can induce vection and motion sickness, particularly during conditions with unstable vestibular feedback. These results have implications for the design of immersive technologies where enhancing or minimizing illusory self-motion is critical for user experience.

Author e-mail: carsoncrewssmith@nevada.unr.edu

Back to Poster Session

Clinical/Hearing Loss (4-7)

4 Auditory sensitivity in migraine – What is causing it?

Sarah M. Haigh*	University of Nevada, Reno
Carson Smith	University of Nevada, Reno

Diagnosis of migraine includes a sensitivity to sound (phonophobia) and we have preliminary data showing that individuals with migraine are more likely to experience hyperacusis (N=132) and misophonia (N=13) compared to headache-free individuals (N=96 and N=8 respectively). We also found that hyperacusis tends to increase in severity 24-hours before the onset of a migraine attack, be most severe during a migraine attack, and reduce 24-hours after the attack has resolved. However, there is no consensus as to the neural mechanisms underlying the auditory sensitivity. We recorded auditory brainstem responses (ABRs) and steady state auditory evoked potentials (SSAEPs) from primary auditory cortex in 28 migraine and 25 headache-free individuals. Auditory clicks were presented at 4Hz over insert earphones for 16mins while ABR was recorded. As the SSAEPs required more time to generate a stable response, the same clicks were presented at 4Hz for 2 seconds 40 times to create a steady-state response. Recordings were collected in the same experimental session. Both ABR and SSAEPs were similar between migraine and headache-free individuals in their magnitude and timing, suggesting that the auditory hypersensitivity in migraine cannot be explained by abnormal functioning in the basilar membrane, the brainstem, or primary auditory cortex. This analysis does exclude the mid latency responses from subcortical areas such as the thalamus. It is also possible that the stimulus (clicks) did not elicit a hypersensitive response and so using click sounds did not reveal which part(s) of the auditory pathway were perturbed when processing uncomfortable sounds. Future explorations may benefit from using more naturalistic everyday sounds that elicit discomfort to understand the neural underpinning of auditory sensitivity in migraine.

Author e-mail: shaigh@unr.edu

Back to Poster Session

5 The effect of underfitted hearing aids on neurophysiologic and behavioral outcomes

Kayla Cormier*	University of Colorado, Boulder
Carly Schimmel	University of Colorado, Boulder
Jack Preheim	University of Colorado, Boulder
Sarah Conrad	University of Colorado, Boulder
Vinaya Manchaiah	University of Colorado, Anschutz Med. Campus
Anu Sharma	University of Colorado, Boulder

Our laboratory is examining the impact of hearing aid treatments, such as hearing aids and cochlear implants using brain based and behavioral outcomes. While previous studies have demonstrated neurocognitive benefits following the use of properly fitted hearing aids, a clinical reality is that many patients hearing aids are under fitted relative to prescriptive hearing aid targets. The aim of this study was to examine the effect of underfitted hearing aids on neurophysiologic and behavioral outcomes. Participants included fifteen older adults with mild to moderate hearing loss whose hearing aids were underfitted relative to NAL-NL2 prescriptive targets. Longitudinal data was collected one-, three-, six- and twelve-months post hearing aid fitting. The cortical auditory P300 event related potential was recorded using high density EEG. Behavioral measures included the Montreal Cognitive Assessment (MoCA) which is a cognitive screening tool, the QuickSIN to measure speech-in-noise abilities, and the Abbreviated Profile of Hearing Aid Benefit (APHAB) to assess subjective hearing aid benefit amongst other behavioral and subjective measures of hearing aid performance. Preliminary results showed an interesting pattern consisting of initial improvement followed by a return to baseline by one year in auditory cortical responses (as indexed by the P300), cognition (as indexed by the MoCA), speech-in-noise abilities (as indexed by the QuickSIN) and self-reported hearing aid benefit (as indexed by the APHAB).In this preliminary dataset, our results may be suggestive of an initial possible placebo effect with underfitted hearing aids followed by a general lack of long-term benefit with underfit hearing aid use.

Author e-mail: kayla.cormier@colorado.edu

Back to Poster Session

6 Detection of sounds and auditory discomfort

Sidney K. Hulsey* University of Nevada, Reno
Sarah M. Haigh University of Nevada, Reno

What makes a sound uncomfortable? Outside the context of music, little is known about the physical parameters of a sound pleasant or uncomfortable. Recent findings have identified some of the basic parameters of auditory stimuli that drive discomfort such as carrier frequency (pitch) and modulation frequency (wobble). The next question is what makes these sounds uncomfortable? We examined if uncomfortable sounds are easier to detect than comfortable sounds. In the first task, we measured the ability to detect amplitude-modulated sounds (wobbles) in noise when manipulating the modulation frequency (2, 8, and 32Hz) and the modulation depth (60% and 100%). In the second task, the participant was presented with the same amplitude-modulated sounds and was asked to rate the discomfort on a scale of 1 (comfortable) to 9 (uncomfortable). Our preliminary data (N=13) indicated that 100% modulation depth was easier to detect amongst the noise (p=.021) and more uncomfortable than the 60% depth (p=.012). The effect of modulation depth on discomfort was most evident for the 32Hz sounds (p=.012). At present, this indicates that the auditory system is particularly sensitive to (can easily detect) certain sounds, and the increased sensitivity can manifest as discomfort.

Author e-mail: sidhskh007@gmail.com

Back to Poster Session

7 Can practice listening with a simulated hearing protection device benefit worker performance while wearing hearing protection?

Patricia McClain* Kansas State University & Children's Mercy Hosp.
Matthew G. Wisniewski Kansas State University

Despite the substantial risk of occupational hearing loss (HL) associated with working in loud environments, many individuals who work in such settings choose not to wear hearing protection devices (HPDs). These individuals often cite decreased work performance as a primary reason for opting out of HPD use. While previous research has been successful in characterizing the overwhelming prevalence of this issue, the question of how to go about improving HPD use still remains. In this study, we investigated the possibility that performance while wearing an HPD could improve if people are given listening training with a simulated HPD. Participants listened to podcasts in which the audio was filtered to simulate wearing an HPD (trained condition, n = 24) or left unfiltered (untrained condition, n = 23). To quantify the impacts of training on speech comprehension and multitasking performance, we employed a recently developed task that involves tracking on-screen moving objects with a computer mouse. In this task, listeners gain hypothetical dollars by closely tracking a target object that is continually updated via speech commands presented under simulated HPD or no HPD conditions. Background noise masked the speech at signal-to-noise ratios of -6 dB to 0 dB throughout. Results suggest that listening training can have positive impacts on performance. The trained condition outperformed the untrained condition in target object tracking while listening with and without a simulated HPD. Though this is a small study, with an artificial work task, our data suggest that the development and evaluation of listening training programs is worthy of further study. We believe that incorporating listening training into hearing loss prevention programs in the workplace could serve to improve HPD use rates by addressing workers' core performance complaints.

Author e-mail: pattymcclain88@gmail.com

Back to Poster Session

Crossmodal Perception (8-10)

8 “Synekinian pairs”: Manual-vocal gesture integration in experimental contexts

Grégory Beller*	Hochschule für Musik & Theater, Hamburg
Georg Hajdu	Hochschule für Musik & Theater, Hamburg

This study investigates the intuitive relationship between manual gestures and vocal syllables, grounded in shared neurobiological and motor systems. Participants were tasked with producing "Synekinian pairs," consisting of vocal syllables paired with instinctive gestures. Through qualitative and quantitative analyses, the findings reveal consistent patterns in manual-vocal associations, offering insights into gesture-based sound design, vocal synthesis, gesture notation and human-computer interaction. Future work aims to refine the experimental methodology, expand datasets, and validate these observations across diverse cultural and linguistic contexts.

Author e-mail: gregory.beller@hfmt-hamburg.de

Back to Poster Session

9 Multisensory integration in speech perception

Paméla Trudeau-Fisette* Université de Montréal
Lucie Ménard Université du Québec à Montréal

Multisensory integration (MSI) refers to the ability to assimilate information from multiple modalities (Molholm et al., 2002; Stein et al., 1996; Stein & Meredith, 1993) and plays a key role in the development of speech (Ito et al., 2009; Lametti et al., 2012; Perrier, 1995; Skipper et al., 2007; S. Tremblay et al., 2003). Although auditory information generally occupies a predominant place in speech abilities, we now know that the weight given to auditory, proprioceptive and visual inputs varies from speaker to speaker and that the sensory systems function, to some extent, in a complementary manner (Katseff et al., 2012; Lametti et al., 2012).The aim of the present project is to observe the transfer and complementarity of the sensory systems involved in speech perception and production mechanisms. To achieve this, acoustic and perceptual data from auditory, proprioceptive and visual feedback manipulation tasks performed by 5- to 6-year-old francophone speakers and adults were analyzed. Preliminary results demonstrate that the complementarity of sensory systems is greater in children than in adults, and that there is a significant cross-over between the visual and proprioceptive systems, particularly in young speakers.

Author e-mail: pamela.trudeau-fisette@umontreal.ca

Back to Poster Session

10 Does endogenous visual cueing influence audiovisual integration?

Madison Prudencio*	University of New Brunswick, Saint John
Abigail Daley	University of New Brunswick, Saint John
Jonathan Wilbiks	University of New Brunswick, Saint John

Audiovisual integration capacity (AVIC) allows individuals to navigate their environment by being able to perceive coherent multisensory events. While integration is known to depend on spatial, temporal and semantic congruency, debate remains on the role of attention. Some studies suggest that integration can occur pre-attentively and is limited to one item (Van der Burg et al., 2013), while other studies highlight how selective attention influences integration, particularly under high cognitive load (Alsius & Soto-Faraco, 2011, Fernandez et al., 2015) and moreover other studies demonstrate flexibility under reduced load or congruent stimuli (Wilbiks & Dyson, 2018). Attention—endogenous or exogenous—may modulate this integration, though findings are mixed. The current work examines the effect of endogenous attentional cues on audiovisual integration capacity. Participants completed an audiovisual integration task with a set of 8 visual stimuli (dots) arranged in a circle. These dots were initially set randomly to be either black or white then, a subset of the dots changed polarity rapidly. After a brief retention period, one dot location was probed by means of a red dot, and participants were asked to respond as to whether that location changed at the same time as the tone or not. The validity of the cue was manipulated during the trials independently of the validity of the probe dot. The probe dot was always valid on 50% of the trials then the validity of the cue was necessarily manipulated across blocks, with 75% (high validity), 50% (low validity), or no cues. The results showed that there were no significant differences between cueing conditions ($p = .180$, $\eta p^2 = .102$), suggesting that endogenous cueing has no effect on audiovisual integration; future studies should examine the effect of validity rate of the cue being presented.

Author e-mail: madison.prudencio@unb.ca

Back to Poster Session

11 Quantitatively exploring the relationship between self-reported auditory affective experiences, real-time emotional reactions, and music perception

David J. Vollweiler* University of Nevada, Las Vegas
Rose De Kock University of Nevada, Las Vegas
Joel S. Snyder University of Nevada, Las Vegas
Erin E. Hannon University of Nevada, Las Vegas

Misophonia, musical chills, and autonomous sensory meridian response (ASMR) are surprisingly common, robust affective experiences that arise in response to rich audiovisual stimuli. Although recent evidence suggests overlap in these phenomena in adulthood, relatively little is known about these experiences during childhood and their potential links with developing musicality. Here we ask how these affective experiences are related to one another and to musicality across development. Children (N = 241) and adults (N = 246) completed assessments of misophonia symptom severity, and performed perceptual tasks assessing perception of tonality, musical emotion, beat and meter, and prosody. We also measured real-time emotional reactions, valence, and arousal ratings to videos intended to evoke misophonia, chills, and ASMR. A principal component analysis revealed three components: one for real-time reactions, one for music perception, and one for self-reported misophonia symptom severity. After controlling for age, small correlations using principal component scores were observed between music perception tasks and real-time reactions, and between real-time reactions and retrospective self-reports of misophonia. These findings suggest that retrospective self-reports measure a different—yet slightly related—construct than real-time reactions. Both constructs may reflect important aspects of auditory affective experience that may be useful for understanding the general population as well as those diagnosed with misophonia. Additionally, these findings suggest that those with better music perception skills may also be more likely to experience misophonia, musical chills, and ASMR in real time, implicating potential overlap between aversive and pleasant auditory affective experiences.

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

Back to Poster Session

12 Starting on the right note: Exploring a working memory-centered ear-training pedagogy

Erica R. Knowles*	Berklee College of Music
Leslie Anne Harrison	Berklee College of Music
Bethanie Liu	Berklee College of Music
Mi-Lan Hoang	Berklee College of Music
Ethan Bessette	Berklee College of Music
Logan Bautista	Berklee College of Music
Sarah Naqvi	Berklee College of Music
Noah Solomon	Berklee College of Music

Music literacy is considered an important aspect of competent musicianship, and developing both auditory and written music literacy is a major component of undergraduate Ear Training (ET) courses. Unlike many college-level music programs in the U.S., contemporary/popular music schools, like Berklee College of Music, see a highly heterogeneous group of students. The daunting task for contemporary/popular music educators is how to develop music literacy skills in students with diverse musical backgrounds. In the current research, we identified two major groups of students who enter into our introductory ET courses: visually-trained musicians who need to develop auditory literacy, and aurally-trained musicians who need to develop written music literacy. There is a paucity of research considering the acquisition of ear training skills when coming from aural versus visual music backgrounds, and an incomplete understanding of best pedagogical practices to support student skill development. Considering prior research, two components were identified as being critically important to ET success for both aural and visual musicians: a) musical memory (remembering), and b) schematic knowledge (understanding). Over the course of one year, the current study implemented pedagogical techniques developed to target these areas, and considered the impact on working memory, schematic knowledge, and course-related skills including melodic dictation.

Author e-mail: eknowles@berklee.edu

Back to Poster Session

13 New music consumption habits reshape autobiographical memory to reminiscence tiers

Justine Canio*	University of San Diego
Mia Lopez	Lourdes University
Laura Getz	University of San Diego

Our research explored the interaction of music, personality, and memory building from two key findings in previous research: that personality traits predict music preferences (Rentfrow & Gosling, 2003), and that people recall more vivid autobiographical memories for music from adolescence and early adulthood than other ages (Krumhansl & Zupnick, 2013). However, today's listening environment, shaped by streaming platforms, on-demand access to music, and personalized recommendations, may alter these previous patterns. The current study tested whether these classic findings replicate in today's media context and whether new consumption patterns have surfaced. Undergraduates completed the Short Test of Music Preferences (STOMP), the Ten-Item Personality Inventory (TIPI), and a recognition survey of the Billboard Year-End top two singles from each year 1970-2024. For each five-year block, participants heard five-second excerpts of each song and rated songs on recognition, personal memories, enjoyment, quality, and emotional response.

Instead of a traditional reminiscence bump, we found reminiscence tiers: low personal memories for most music predating parents' birth years, moderate memories for most music between parents' and participants' birth years, and high memories and recognition for music postdating participants' birth years. In terms of the personality and music preferences data, we found no correlations between the STOMP categories and Big 5 personality traits.

A possible explanation for the new results is modern music consumption. While radio, peer culture, and family once shaped music exposure, streaming has transformed how listeners access and consume music. Algorithmic platforms may weaken the reminiscence bump and instead develop "reminiscence tiers". These findings show earlier associations between personality, music preference and memories may not apply in the streaming era. This shift has potential applications in music therapy, digital health, and personalized listening algorithms, highlighting the need to study how emerging streaming platforms shape the connections between music, memory, and identity.

Author e-mail: jurielle9@gmail.com

Back to Poster Session

14 Evaluating potential influences of background music on working memory capacity

Ange-Dominique A. Akesse	James Madison University
Carter Peery*	James Madison University
Michael D. Hall	James Madison University

Listening to background music has been argued to be either beneficial or detrimental to working memory tasks, depending upon task/stimulus conditions (e.g., see Lehmann & Seufert, 2017). This raises the possibility that characteristics of the music could impact working memory capacity, including in learning situations. The current investigation sought to determine conditions where relatively unfamiliar instrumental music might impact a common measure of working memory capacity, the Backwards Digit Span Task (BDST), across two experiments with undergraduate listeners.

Experiment 1 evaluated potential effects of music genre on BDST span (4-9 digits) by counterbalancing classical, jazz, and progressive rock excerpts relative to a control condition (rain recordings). Higher spans were obtained across all music conditions relative to the control, reaching statistical significance for classical music. Furthermore, mean spans improved as the average spectral centroid of the sounds lowered, suggesting reduced demands on attention. Listeners also indicated that the rain sounds in the control condition were often distracting, which questions whether the background music indeed benefited working memory spans.

Experiment 2 addressed this concern by repeating the task with a silent control and extended genre considerations by substituting lesser-known pop music for the rock samples. It also included ratings of song familiarity and listening preferences following each musical sample to see if either variable could predict performance. While traditional BDST scoring revealed no significant differences across background listening conditions, summed scores revealed a marginal decline in spans for the control. These findings were predicted by neither familiarity with the music nor listening preferences.

The results across experiments suggest that background music can slightly benefit working memory span, particularly if the music does not contain much high-frequency content, and thus demands minimal attention. Potential limitations of standard scoring for the BDST, along with concerns for engagement in the task, will be discussed.

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

Back to Poster Session

Speech & Language (15-21)

15 Transfer of statistical learning from speech perception to production generalizes to reading

Kyle Huffaker*	Indiana University, Bloomington
Lori L. Holt	University of Texas at Austin
Nazbanou Nozari	Indiana University, Bloomington

Exposure to speech acoustics that subtly depart from language norms shifts perception, with transfer to speech production. However, all studies reporting such transfer have used auditory repetition tasks. Therefore, it is unclear whether perception-production transfer in the acoustic-phonetic domain extends to tasks without an auditory model to probe production. We answer this question in three experiments. Experiment 1 shows that people read aloud the words BEER and PEER differently after exposure to auditory samples of 'beer' and 'peer' drawn from a distribution of standard American English vs. a distribution of slightly accented speech. Experiments 2 and 3 replicate this finding and show generalization to reading a new word pair (BEACH/PEACH) and a new nonword pair (BEETH/PEETH). Collectively, these results demonstrate that the perception-production transfer in the acoustic-phonetic domain extends beyond auditory repetition tasks to production tasks without an explicit auditory model, and that this transfer generalizes to new syllables.

Author e-mail: kydahuff@iu.edu

Back to Poster Session

16 CNV as a neural signature of speech production modulated by complexity and planning

Mohsen Parsa* Arizona State University

Speech production requires precise coordination between cognitive planning and motor execution. The Contingent Negative Variation (CNV), a slow cortical potential that emerges between a warning cue and an imperative stimulus, is thought to reflect anticipatory attention, temporal preparation, and motor readiness. While CNV has been extensively studied in motor control, its role in speech planning remains less clear. In this study, we investigated how CNV amplitude is modulated by task complexity and available preparation time during overt and covert speech.

EEG was recorded from 20 healthy adults as they completed a cued word production paradigm. On each trial, a visually presented word (monosyllabic or multisyllabic) appeared in white font and later changed color, serving as the go-signal. Participants either produced the word aloud (speaking condition) or read it silently (silent reading condition). Planning demands were manipulated by varying the interval between stimulus onset and the go-signal (short vs. long).

Preliminary analyses revealed CNV preceding speech onset, maximal over fronto-central electrodes. Importantly, CNV amplitude was greater in long preparation trials compared to short ones, suggesting increased anticipatory activity when participants had more time to prepare. Moreover, multisyllabic words elicited larger CNVs than monosyllabic words, consistent with higher planning demands. Speaking trials showed stronger CNVs than silent reading, reflecting additional motor preparation.

These findings provide novel evidence that CNV is sensitive to both complexity and temporal context in speech production. By demonstrating that CNV reflects preparatory mechanisms at the interface of cognitive-linguistic planning and motor execution, this work contributes to understanding how the brain organizes speech in real time. Future comparisons with clinical populations, such as individuals who stutter, may clarify whether atypical CNV dynamics underlie speech production difficulties.

Author e-mail: mparsa1@asu.edu

Back to Poster Session

17 Cognitive control during speech recognition in noise in ADHD

Victoria Holcomb* Auburn University
Susan Teubner-Rhodes Auburn University

Attention-Deficit/Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder characterized by inattention, hyperactivity, and impulsivity. ADHD can impact higher-order cognitive processes, including speech recognition and decision making. Individuals with ADHD exhibit decreased accuracy in speech recognition in noise and display higher impulsivity linked to poor cognitive control. Recognizing speech in noisy environments may involve cognitive control to filter out irrelevant auditory information and select amongst activated phonological competitors. However, it is unknown how ADHD affects the application of cognitive control during speech recognition in noise. Our study aims to understand how ADHD influences the deployment of cognitive control during speech recognition in noise, using a picture-speech recognition task designed to trigger cognitive control through phonological conflict.

Native-English speaking participants ages 18-35 with or without a formal diagnosis of ADHD were recruited. They had normal or corrected-to-normal (20/40) vision, hearing loss < 25 dB, and no history of neurological disorders. All participants completed demographic and ADHD symptomology questionaries as well as screenings for vision and hearing. Eligible participants completed a speech-in-noise recognition task in which target words were presented at +6 dB signal-to-noise ratio against 60 dB SPL multi-talker babble alongside congruent or incongruent photorealistic pictures of objects. On congruent trials, the picture matched the spoken word. On incongruent trials, the picture was a phonological neighbor of the spoken word to induce conflict. Increased speech recognition accuracy following conflict indexes cognitive control-related improvements.

Preliminary results (n = 8) showed that participants with ADHD exhibited a significant improvement in word recognition following conflict (+16%-points; p = .02) while healthy controls did not (+4%-points, p = .76). This pattern reflected decreased accuracy (-11%-points) on initial conflict trials in participants with ADHD compared to controls. Results suggest that young adults with ADHD apply cognitive control reactively rather than proactively when recognizing speech in noisy environments.

Author e-mail: vjh0006@auburn.edu

Back to Poster Session

18 Generalization of statistical learning of vowels in speech perception and production

Lin Zhou*	University of Texas at Austin
Craig Thorburn	University of Texas at Austin
Nazbanou Nozari	University of Texas at Austin
Lori L. Holt	University of Texas at Austin

The regularities of one's native language are encoded in long-term representations that shape how speech is perceived. When incoming acoustic signals deviate from these regularities—for example, due to foreign accents or regional dialects—listeners rapidly adapt, even during passive exposure to another talker's speech. Prior studies suggest that such perceptual shifts can also influence listeners' own speech production. However, this work has focused on consonants. Therefore, it is unclear if transfer also affects vowels. Given that vowels carry important information about accent and dialect, the current study aimed to fill this critical gap. To this end, we manipulated the relationship between spectral energy and vowel duration in the /ɛ/-/æ/ contrast. We passively exposed participants to speech that either conformed to English norms or diverged in an accent-like manner, by perturbing the relationship between spectral energy and duration. Native American English speakers normally rely on both spectral energy and vowel duration to categorize /ɛ/-/æ/ . When exposed to the accent, listeners rapidly down-weighted reliance on duration. Importantly, this perceptual reweighting also transferred to production, with duration differentiating vowels less in the context of the accent. The perceptual influence generalized to non-lexical items never heard in the accent and transferred to impact production. Together with past findings, these results demonstrate that short-term statistical manipulations of the speech environment not only retune internal perceptual mappings but also generalize robustly to production across both consonants and vowels, with generalization to new contexts.

Author e-mail: lin.zhou@austin.utexas.edu

Back to Poster Session

19 Exploring decision-making processes in speech production

Jihee Renee Kang* Arizona State University
Ayoub Daliri Arizona State University

Decision-making has been a topic of interest in various fields. Among a range of topics, decision-making under uncertainty and time pressure has significant implications in daily lives. Previous studies investigated people's performance tendency and decision-making processes on the task where they are asked to change their movement plans suddenly (Chikazoe et al., 2009, Contemori & Carroll, 2025, Swainson et al., 2003).

These studies were limited to motor movement, and work in the speech domain remains underdeveloped. Therefore, the current study looked at the topic in the field of speech. The purpose of this study is to gain an understanding of how people would decide to react when a sudden change of speech planning is required. Would they ignore the suddenly presented change and produce a response using the initially created speech plan, or would they decide to build a new speech plan? During these processes, would their speech show different characteristics compared to those responses when the target words did not change suddenly?

We collected speech responses in two types: Non-switch and Switch trials. In the Non-switch trial, the target word that participants had to produce did not change. In the Switch trials, the ultimate target word suddenly changed and thus was different from the initially presented word, so participants were required to quickly change their speech plan and produce a new target word with their new speech plan. We measured and analyzed participants' reaction time in each of the experiment conditions, error rate in 8 distinct categories, and formants of the vowels in their responses.

Our preliminary results show that participants tended to make errors in the Switch trials, with different types of error categories. Reaction time decreased as the experiment proceeded, as the participants completed subsequent blocks. Formants of the vowels demonstrated differences between those in the Switch and Non-switch conditions.

Author e-mail: jkang94@asu.edu

Back to Poster Session

20 Direct brain recordings show strengthening prioritization of speech over music

Rajvi Agravat*	University of Texas at Austin
Maansu Desai	University of Texas at Austin
Alyssa M. Field	University of Texas at Austin
Gabrielle Foox	Texas Children's Hospital, Baylor Col. of Medicine
Sandra Georges	Texas Children's Hospital, Baylor Col. of Medicine
Jacob Leisawitz	Texas Children's Hospital, Baylor Col. of Medicine
Saman Asghar	Texas Children's Hospital, Baylor Col. of Medicine
Anne E. Anderson	Texas Children's Hospital, Baylor Col. of Medicine
Dave Clarke	Dell Children's Medical Center
Elizabeth C. Tyler-Kabara	Dell Children's Medical Center
Andrew J. Watrous	Texas Children's Hospital, Baylor Col. of Medicine
Howard L. Weiner	Texas Children's Hospital, Baylor Col. of Medicine
Liberty S. Hamilton	UT Austin, UC Berkeley, UC San Francisco

Our brains are constantly filtering sounds. Are some temporal lobe areas more selective to certain sound streams? While extensively studied in adults, understanding how auditory stream selectivity develops and matures throughout childhood remains a critical gap in our knowledge. Using intracranial stereo-electroencephalography (sEEG), we investigated this implicit auditory attention question by presenting 51 pediatric participants (ages 4-22, 29M/23F, over 8000 electrodes total, with over 1000 temporal lobe electrodes total) with audiovisual movie trailers that contained both speech and music. Our naturalistic approach provides critical insights into real-world auditory processing beyond traditional laboratory paradigms. We extracted and analyzed high gamma band activity (70-150 Hz) to index local neural firing. Data were analyzed using a computational source-separated STRF analysis. We used deep neural networks to decompose the mixed speech-music stimuli into isolated speech and music components, then built separate encoding models for each condition (speech only, music only, and original speech-music mixture). Although the participants heard only the original speech-music mixture, statistical analyses using linear mixed effect models found that in superior temporal gyrus (STG), the speech only model significantly outperformed the mixed and music only models ($p<0.0004$ for both comparisons). Analysis of age-related effects showed a highly significant relationship between development stage and brain region ($p<0.0001$), with STG showing the strongest developmental change in speech selectivity, indicating that the pediatric brain demonstrates an intrinsic neural bias toward socially relevant auditory information that emerges early and strengthens throughout development. Our findings demonstrate how different brain regions implicitly attend to specific components of mixed auditory streams, with specialized temporal lobe networks enabling the separation and processing of speech and music simultaneously, even in the absence of explicit attentional directives. This suggests that selective attention processing in implicit attentional scenarios may be driven largely by bottom-up mechanisms, automatic, stimulus-driven processes, rather than a purely conscious, top-down mechanism. These results have significant implications for understanding

developmental communication disorders and may inform targeted therapeutic interventions for children with auditory processing deficits.

Author e-mail: rajvi@utexas.edu

Back to Poster Session

21 Speech envelope modulation detection across frequency bands

Milad Yousefi*	University of Illinois, Urbana-Champaign
Daniel Fogerty	University of Illinois, Urbana-Champaign

Slow amplitude modulations (<16 Hz) are crucial for prosody, stress, and speech intelligibility. The ability of a listener to detect these modulations is important for speech perception. However, standard psychoacoustic tests typically use simple sinusoidal modulators rather than the complex envelopes of real speech, potentially limiting ecological validity. This study introduced a speech-derived envelope approach to measure modulation detection thresholds for two frequency bands. Twenty young adult listeners with normal hearing (ages 18–30) participated in a three-interval, two-alternative forced-choice task to identify which 500-ms speech-shaped noise burst was amplitude-modulated by a speech envelope (vs. unmodulated). The noise stimuli were filtered into two one-third-octave bands: a lower band centered at 500 Hz and a higher band centered at 3150 Hz. On each trial, only one band carried the speech-derived modulation while the other remained unmodulated, isolating detection to the target region without across-band interference. An adaptive tracking procedure, Parameter Estimation by Sequential Testing (PEST), adjusted the modulation depth to estimate the 50% modulation detection threshold for each band.

Results showed that modulation detection was significantly better at 3150 Hz than at 500 Hz, with listeners able to tolerate greater compression of the amplitude modulation in that band. This potentially indicates greater sensitivity to speech envelope cues at higher audio frequencies, possibly reflecting the prominence of consonant-related modulations in that region. By using natural speech envelopes instead of synthetic modulators, this paradigm provides a more realistic measure of speech temporal processing across frequency bands. This method shows promise for research and clinical applications. For example, this method can be used to examine how aging or hearing loss may affect the perception of speech rhythm across different frequency regions. [Work supported by NIH/NIDCD]

Author e-mail: milady2@illinois.edu

Back to Poster Session