# *APCAM 2016*

## 15th Annual Auditory Perception, Cognition, and Action Meeting

***Thursday, November 17th***

***Sheraton Boston Hotel***
***Back Bay C (2nd Floor)***
***Boston, Massachusetts***

*Program sponsored by*

WASHBURN UNIVERSITY

## Welcome to APCAM 2016

Welcome to the fifteenth annual Auditory Perception, Cognition, and Action Meeting (APCAM 2016) at the Sheraton Boston Hotel in Boston, Massachusetts. This means that for fifteen years APCAM has continued to deliver on its mission ". . . to bring together researchers from various theoretical perspectives to present focused research on auditory cognition, perception, and aurally guided action". If we are to continue to advance auditory science, then we all need outlets to collectively present both basic and applied research that involves diverse stimuli and considers multiple levels of processing without reliance on a singular theoretical model. We believe APCAM to be one of the only conferences that combines all of these within a single session.

This past year has been a big one for the meeting and its attendees. A special issue of *Journal of Cognitive Psychology* is being finalized that highlights speech- and language-related work from a previous APCAM. Following prompts from last year's meeting, we surveyed attendees regarding how to improve the conference website, as well as interest in pursuing development of a journal that reflects APCAM's themes. Thank you to the large number of respondents for all the useful information you provided. In response to feedback, the conference website (www.apcam.us) has been expanded to include a page that provides descriptions of, and links to, the laboratory sites of participating attendees. This resource should launch right around the time of our meeting, and we would very much like to see this grow in the future in order to maximize its utility. If you, as an APCAM attendee, would like to include your lab's information on the site, then simply use the types and order of information provided by other laboratories on the site as a template. You can send your information to Mike Russell at mike.russell@washburn.edu. Additionally, we have been looking into the possibility of streaming oral presentations during this and future meetings. Finally, the survey indicated that the vast majority of respondents were very supportive of the possibility of a new journal that is consistent with APCAM's mission. We have therefore proposed such a new journal (currently under review) to Taylor & Francis. We will be sure to update you with any further information regarding this proposal as it becomes available.

There are signs that the meeting has started to sustain gradual growth. Work presented at the Boston meeting will reflect 48 submissions from several countries/continents. An overview of abstracts reveals emphases in both general and domain-specific (e.g., music, speech, and language, etc.) areas of perception, action, and attention. The program also includes two distinguished invited speakers and a panel discussion concerning an ongoing research topic/problem. Over the years, we have largely depended on word of mouth to spread the word about our meeting, and it is clear that the program benefits from new attendees that were encouraged to attend by APCAM regulars. In fact, the infusion of new blood has traditionally had long-term consequences, as first-time attendees often become APCAM regulars. We are dependent upon you to make sure that this trend continues. We therefore kindly ask that if you enjoy your APCAM experience, then please pass that information along to other colleagues and bring them to our next meeting (to take place in Vancouver, British Columbia, Canada on November 9, 2017). This will ensure outstanding APCAM programs for years to come.

This year's APCAM has been made possible by important continuing contributions from a couple of organizations. First, the Psychonomic Society again agreed to cover all expenses associated with satellite meeting rooms, A/V equipment, and poster displays. As a result, our meeting continues to be offered as a free event. Please be sure to support the subsequent meeting of the Psychonomic Society and to indicate your appreciation of their continued support of satellite meetings. Likewise, for the past several years, Washburn University has covered all money and time costs associated with producing conference programs. The Organizing Committee's Mike Russell has continuously overseen this effort, so please be sure to share your gratitude for his hard work and the support of his department/institution.

Ultimately, heartfelt thanks also go to you—the authors for choosing to share your compelling work at this meeting, and attendees for feeding active discussion concerning that work and potential directions for future research. Thanks to you, we believe that this year's APCAM promises consistently interesting and thought-provoking sessions. We hope that you concur. As always, if there are issues that arise during the meeting, or thoughts that you have for further enriching the program or its execution in the future, then please do not hesitate to share those concerns or thoughts with any of us on the committee at any point. We welcome your feedback. Have a pleasant and productive day at APCAM.

Sincerely,
The APCAM 2016 Organizing Committee

| | | |
|---|---|---|
| Michael D. Hall (Chair) | Devin McAuley | John Neuhoff (Founder) |
| Kristopher (Jake) Patten | Peter Q. Pfordresher | Mike Russell (Assistant Chair) |

| APCAM 2016 Schedule | | |
|---|---|---|
| **8:00** | **Registration – Back Bay C, 2nd Floor, Sheraton Boston** | |
| **8:30** | **Opening Remarks** | |
| | ***Music (abstracts pages 9 – 11)*** | |
| **8:45** | Investigating tempo biases in the mental representation of rhythm using the method of serial reproduction | **Elisa Fromboluti*<br>J. Devin McAuley** |
| **9:00** | Effect of variable tempo learning on skill acquisition | **Baptiste Caramiaux*<br>Marcelo Wanderley<br>Caroline Palmer** |
| **9:15** | Perceptual ratings of rhythmic analogs of musical pitch intervals and chords | **Aniruddh Patel*<br>Alek Razdan** |
| **9:30** | Effects of pitch distance and harmonic similarity on recognition of transposed melodies | **Abigail Kleinsmith*<br>W. Neill** |
| **9:45** | Progress from analytic to global perception of modulations with increased familiarity with music | **Walter Dowling*<br>Rachna Raman** |
| **10:00** | Statistical considerations for evaluating the influence of musical training | **Heather Daly*<br>Michael Hall** |
| **10:15** | **Break (15 mins)** | |

| | | |
|---|---|---|
| ***Invited Address (abstract page 12)*** | | |
| **10:30** | ***Cortical control of auditory attention*** | ***Barbara Shinn-Cunningham*** |
| ***General (abstracts page 13 – 14)*** | | |
| **11:00** | The effects of action and attention on auditory learning | **Elizabeth Casserly*<br>Hunter Drews,** |
| **11:15** | The effect of auditory cues on co-modulation masking release: The role of attention in the detection of sounds masked by narrow-band noise | **Doug Alards-Tomalin*<br>Todd Mondor<br>Jason Leboe-McGowan** |
| **11:30** | The impact of working memory load and noise on the online processing of spoken words: Evidence from eye movements | **Boaz Ben-David*<br>gal Nitsan<br>Daniel Algom<br>Arthur Wingfield** |
| **11:45** | Spatial release from informational masking modulates auditory evoked potentials | **Benjamin Zobel*<br>Lisa Sanders<br>Richard Freyman** |
| ***Poster Session (12 noon – 1:20 PM)<br>Ballroom A – C (located on Level 3)<br>Abstracts located on pages 21 –35*** | | |
| ***Lunch (1:20 – 2:00 PM)*** | | |
| ***Invited Address (abstract page 16)*** | | |
| **2:00** | ***A Revised Duplex Model for Sound Localization*** | ***William M. Hartmann*** |
| ***Panel Discussion (abstract pages 17)*** | | |
| **2:30<br>to<br>3:15** | ***Unifying principles and individual differences: What should we focus on to understand speech?*** | **Joseph Toscano*<br>Bob McMurray<br>Emily Myers<br>Matt Goldrick** |

| | | |
|---|---|---|
| **3:15** | **Break (15 minutes)** | |
| | ***Perception & Action (abstracts pages 18 – 20)*** | |
| **3:30** | Looming sounds are perceived as faster than receding sounds | **John Neuhoff*<br>Christian Golden** |
| **3:45** | Hearing gestures and drawing sounds: Auditory and multisensory perception of biological movements | **Etienne Thoret*<br>Mitsuko Aramaki<br>Sølvi Ystad<br>Richard Kronland-Martinet** |
| **4:00** | Music expertise alters the perception of auditory objects | **Mrinmayi Kulkarni*<br>Katherine Potkonjak<br>Richard Randall<br>Adam Greenberg** |
| **4:15** | Comparing perception–action coupling in groups of improvising and non-improvising musicians with an EEG Stroop task | **Andrew Goldman*<br>Paul Sajda** |
| **4:30** | Bass drop: Establishing a link between low-frequency tones and movement | **Michael Hove*** |
| **4:45** | A mutual adaptation model of joint action in music performance | **Alexander Demos*<br>Hamed Laeghi<br>Caroline Palmer** |
| **5:00** | **Updates & Closing Remarks** | |

| Posters (abstracts located on pages 21 – 35) | | |
|---|---|---|
| **1** | Subcortical encoding of sound: Interaction between attentional demand and stimulus regularity | **Zilong Xie*<br>Bharath Chandrasekaran** |
| **2** | Developing an auditory and visual cross-modal continuous performance task for evaluating concussion | **Christopher Koch*<br>Taylor Charbonnier<br>Kristin Dissinger<br>Steven Egeberg<br>Matthew Johnson<br>Lindsay Levanen<br>Matthew Scott** |
| **3** | More attention, less imitation? | **Marie Postma*** |
| **4** | Effects of ruminative mind wandering episodes on speech imitation, lexical associations and eye behavior | **Marie Postma*** |
| **5** | The influence of musical instrumentation on the occurrence of audiovisual associative links | **Elena Anisimova*** |
| **6** | The effect of stimulus frequency, quality, duration, and location on temporal order judgment threshold | **Leah Fostick*<br>Harvey Babkoff** |
| **7** | Phonetic category activation can drive rapid speech adaptation | **Yunan Wu*<br>Lori Holt** |
| **8** | Happy mood makes music sound brighter | **Ryan Brigante*<br>Bart Rypma** |
| **9** | Listeners anticipate upcoming vowels before identifying the current fricative | **Kayleen Schreiber*<br>Bob McMurray** |
| **10** | Modeling audiovisual cue integration in speech | **Laura Getz*<br>Elke Nordeen<br>Sarah Vrabic<br>Joseph Toscano** |

| | | |
|---|---|---|
| **11** | Individual differences in dynamic adaptation during lexically-guided perceptual learning | **David Saltzman*** **Emily Myers** |
| **12** | Working memory capacity predicts the pupil dilation response to auditory deviation | **Alexandre Marois*** **Johnathan Crépeau** **François Vachon** |
| **13** | Using sound to detect the presence of an occluding object: Can physics alone explain human perception? | **Stephanie Brown** **Michael K. Russell** |
| **14** | Using echolocation to detect the existence of openings and object location | **Stephanie Brown** **Michael K. Russell** |
| **15** | Examining the optimal and observed bounds of sound class discrimination | **Mattson Ogg*** **Robert Slevc** |
| **16** | Millenials write less-positive things than their parents about recent pop songs | **Anne Gilman*** **Clayton Reichart** |
| **17** | Selective and divided attention in polyphonic melodies: Training related advantages | **George Seror III*** |
| **18** | The influence of inversions on the perception of major and minor chords | **Vincenzo Belli*** **Abigail Kleinsmith** **W. Neill** |
| **19** | Which cues do listeners use? Discovering networks of phonetic cues for speech sound categorization using a graph theoretic approach | **Anne Marie Crinnion*** **Beth Malmskog** **Joseph Toscano** |
| **20** | Measuring working memory: A theoretical exploration of musical working memory | **Juan Ventura*** **Emily Elliott** |
| **21** | The interaction of repetition and short-term memory in melodic dictation tasks | **Nathan Cornelius*** **Jenine Brown** |

| | | |
|---|---|---|
| **22** | Influence of spectrotemporal modulations on the recognition of musical instrument sounds | **Etienne Thoret*<br>Philippe Depalle<br>Stephen McAdams** |
| **23** | Error correction neural correlates of auditory and visual sensorimotor synchronization | **Daniel Comstock*<br>Ramesh Balasubramaniam** |
| **24** | Tonality without structure: Effects of tonal uncertainty and drone tones on emotional response to melodies | **Jeff Bostwick*<br>W. Neill** |
| **25** | The effects of tonal-harmonic structure on visual narrative processing | **Morgan Patrick*<br>Sheila Blumstein<br>Neil Cohn** |
| **26** | Preference for colored audio noise | **Andrew Slifkin*<br>Katrina Wolfgang** |
| **27** | Comparing linguistic prediction in musicians and non-musicians | **Allison Fogel*<br>Eddie Wlotko<br>Gina Kuperberg<br>Aniruddh Patel** |
| **28** | Systematic variation in rhythm production as tempo changes | **Scott D. Barton<br>Laura Getz<br>Michael Kubovy** |
| **29** | How do we decide when an auditory sequence ends and another begins? | **Lauren Amick*<br>Audrey Drotos<br>J. Devin McAuley** |

# Oral Presentations

**8:45**

## Investigating tempo biases in the mental representation of rhythm using the method of serial reproduction

**Elisa Fromboluti*** *Michigan State University*
**J. Devin McAuley** *Michigan State University*

Previous literature has consistently found that adults prefer to produce and listen to sequences at a rate of approximately 100 beats-per-minute (BPM). The present research examined the influence of tempo preferences on the mental representation of auditory and visual rhythms using the method of serial reproduction (SR). In two experiments, SR chains were initiated with isochronous seed rhythms across a broad range of rates. Each participant reproduced the tempo of the stimulus rhythm by tapping, with one participant's reproduction then serving as the stimulus for the next participant in the chain. Reproductions at each step were predicted to systematically distort the to-be- reproduced stimulus towards preferred tempo, thereby revealing a preference for a tempo of approximately 100 BPM by the final reproduction in the chain, regardless of the chain's seed tempo. Consistent with this general prediction, Experiment 1 found that SR chains initiated with auditory or visual seed rhythms ranging from 50 to 200 BPM converged to approximately 120 BPM. This result was replicated in Experiment 2 across a broader range of seed tempi ranging from 35 to 400 BPM. Experiment 2 additionally examined motor contributions to tempo encoding across modalities by having participants either passively observe or actively synchronize with the to-be-reproduced rhythm during exposure. Results from Experiment 2 showed that synchronizing with the rhythm during encoding reduced reproduction bias relative to passively observing the rhythm, but more so for auditory than for visual rhythms. These results contribute to a growing body of evidence that a preference for a tempo of approximately 100 BPM biases the representation of rhythms and suggest that the role of the motor system in temporal encoding differs across auditory and visual modalities. Moreover, the experiments establish the method of serial reproduction as a viable tool for examining biases in the mental representation of rhythm.

Email: ekfromboluti@gmail.com, dmcauley.msu@gmail.com

**9:00**

## Effect of variable tempo learning on skill acquisition

**Baptiste Caramiaux*** *McGill University, IRCAM*
**Marcelo Wanderley** *McGill University*
**Caroline Palmer** *McGill University*

Moving efficiently and at the right time is important in complex human skills such as music performance. The acquisition of such motor skills remains largely unexplored, especially the role of learning schedules in acquisition of timed motor skills. Varied learning schedules are known to facilitate transfer to other motor tasks, and retention of a learned task in spatial movements. Facilitation of transfer of learning through variability in learning schedules has been reported as evidence of structure learning in visuomotor tasks. Variability during practice of wide left-hand interval leaps in piano performance led to better transfer to untrained motor tasks right after training, but not after a retention period. We examined the effects of a variable tempo learning schedule on timing skill acquisition with 24 non-musician participants who learned an 8- note sequence on a piano keyboard. 12 of the participants practiced the sequence at 2 tempi ("2-tempo Group"), while 12 participants practiced the sequence at 6 different tempi ("6-tempo Group"). Both groups performed the same number of total practice trials. We examined timing regularity in three transfer tests: one at a novel (average) tempo, one at a novel fast tempo and one at a previously practiced slow tempo. The 2- tempo Group showed lower variability at the novel (mean) tempo, contrary to previous findings in the spatial domain. The novel (average) tempo yielded the least variability across all transfer conditions, suggesting most transfer from the different tempo practice conditions. The 2-tempo Group also performed with lower variability at the unfamiliar fast tempo, even though the 6-tempo Group practiced fast tempi closer to the novel test tempo. Finally, the 2-tempo Group performed the familiar slow tempo with less variability. We discuss implications of learning schedules on transfer of motor skills in temporal and spatial domains.

Email: baptiste.caramiaux@mail.mcgill.ca, marcelo.wanderley@mcgill.ca, caroline.palmer@mcgill.ca

**9:15**

## Perceptual ratings of rhythmic analogs of musical pitch intervals and chords

**Aniruddh Patel*** *Tufts University*
**Alek Razdan** *Northeastern University*

The current study quantifies human preferences for 'rhythmic intervals' (polyrhythms) formed by expressing the frequency ratios of common Western musical pitch intervals and triads as rhythms. We hypothesized that pleasantness ratings of these rhythmic stimuli would resemble those of their pitch counterparts. Consistent with this hypothesis, we find that pleasantness ratings of rhythmic intervals and triads correlate significantly with pleasantness ratings found in earlier research on corresponding pitch intervals and triads. Pleasantness ratings for our polyrhythms were inversely related to ratings of their rhythmic complexity, and positively related to ratings of how strongly they induced a sense of an underlying beat ('beat induction') and how much they made one want to move ('groove'). Possible mechanisms underlying these results will be discussed.

Email: a.patel@tufts.edu, alekrazdan@hotmail.com

**9:30**

## Effects of pitch distance and harmonic similarity on recognition of transposed melodies

**Abigail Kleinsmith*** *University at Albany, SUNY*
**W. Neill** *University at Albany, SUNY*

What factors influence a person's ability to recognize a melody when it is transposed into a new key? Our previous research (Kleinsmith, 2015) has demonstrated a robust effect of pitch distance on memory for melodies: the more distant in pitch height a given melody is from the key of a previously studied target melody, the more difficult it is to discriminate it from a foil in the same key on a subsequent recognition memory test. We familiarized participants with a target melody in two keys close in pitch height (C and D). Recognition of the target melody was better in a physically closer key (C#) than in a more distant, but more harmonically related, key (G). We have replicated this pitch-distance effect in experiments where participants were familiarized with a target melody in two physically distant keys (C and F#), familiarized with melodies for which it is impossible to establish a key (composed of whole-tones), and tested on a target melody presented in a different pitch class (a test of octave equivalence). When we performed a stronger test of harmonic similarity by holding pitch distance constant across conditions, we failed to find any benefit of harmonic relatedness on recognition performance. Taken together, these results imply that pitch distance is much more important than harmonic distance for recognition of transposed melodies.

Email: akleinsmith@albany.edu, wneill@albany.edu

**9:45**

## Progress from analytic to global perception of modulations with increased familiarity with music

**Walter Dowling*** *University of Texas, Dallas*
**Rachna Raman** *University of Texas, Dallas*

We present three experiments using Toiviainen and Krumhansl's (2003) continuous probe-tone method to track listeners' perception of tonal modulations. Listeners hear music in one ear and a steady probe in the other, continuously rating how well the probe goes with the music. The excerpt repeats 12 times, using each of the 12 possible semitones. We construct tonal hierarchy profiles for 10-s segments throughout the piece, and correlate them with the profiles for the keys involved. Shifts in key are reflected in the pattern of those correlations. In Study 1 Indian and Western music teachers heard South Indian classical songs. Only Indian teachers were familiar with the songs. Surprisingly, differentiation of modes in modulations was clearer for the Westerners. We thought perhaps greater familiarity led to more global perception, in which a piece is heard not as sharply modulating but as blending a cluster of different keys throughout. In Study 2 highly trained Western musicians heard 2-min Haydn quartet excerpts. We compared performance during the first three trials in which they were unfamiliar with the pieces with the last three trials where they were more familiar. Differentiation of the keys was greater at first, and later became more global. In Study 3 we manipulated familiarity strongly with student orchestra members in three sessions: before encountering the piece they were to learn, in the middle of rehearsals, and after performing the piece. However, there was little change in their responses from beginning to end, perhaps because their active involvement maintained their analytic mode. In conclusion, there is some tendency for musicians to hear a piece more and more as a unified whole with increasing familiarity, but not always. In some cases they form a holistic representation of the piece as a cluster of related keys, rather than as a sharply differentiated sequence of individual episodes.

Email: jdowling@utdallas.edu, rachnaraman@hotmail.com

**10:00**

## Statistical considerations for evaluating the influence of musical training

**Heather Daly*** *James Madison University & US Army Research Laboratory*
**Michael Hall** *James Madison University*

Traditionally, studies concerning the influence of musicianship on task performance have at least partially defined musical training as the number of years of training on a musical instrument or voice. Individuals are usually sorted into two groups—"musicians" and "nonmusicians"—based on cut-off scores, even though training is a continuous variable. Categorizing continuous variables introduces several problems, including reduced power, spurious significance, loss of measurement reliability, and elimination of individual differences (MacCallum, Zhang, Preacher, & Rucker, 2002). To examine how common and variable this practice might be, we reviewed a cross- section of (38) pitch studies from the past 25 years. Seventy-one percent of studies relied on years of training, and reflected 20 different cut-offs (2-14 years). The adverse impact of different cut-offs was investigated using two data sets from our laboratory [one from a published study concerning pitch and timbre interaction (Becker & Hall, 2014), and another from a recently completed investigation of change deafness (Daly, Hall, Gaston, & Dickerson, in preparation)], as well as simulated data sets varying in the distribution of musical training (positively skewed v. normal). Cut-offs were systematically adjusted within each data set, and outcomes were compared against each other, as well as against a regression model in which musical training was treated as a continuous variable. As cut-offs varied, analyses of actual data sets revealed substantial variability in effect sizes (e.g., partial eta squared = .067-.345 for pitch data), and only occasional significance; obtained effect sizes varied even more widely within simulated data sets. In contrast, regression analyses consistently revealed significance, with effect sizes that were comparable to, or exceeded, those obtained using cut-off scores. These findings suggest that musicianship based upon years of training should be modeled as a continuous variable to maintain naturally occurring variability while facilitating more appropriate comparisons of findings across studies.

Email: dalyhr@dukes.jmu.edu, hallmd@jmu.edu, jeremy.r.gaston.civ@mail.mil, kelly.l.dickerson3.civ@mail.mil

**10:15 Break (15 mins)**

**10:30**

## *Invited Address*

### *Cortical control of auditory attention*

**Barbara Shinn-Cunningham** *Boston University*

In many social settings, there are multiple, competing sounds vying for attention. The ability to separate sound streams coming from different sources and focus on whatever source you want to understand is critical for communication in such environments. This talk reviews behavioral, EEG, and fMRI studies that explore how listeners control auditory attention. Results show that when listeners decide to focus attention on a sound stream from a particular direction or from a particular talker, there is preparatory activity in various brain networks. Once the sound stimuli begin to play, the cortical representation of the competing sound streams is modulated, such that responses to an attended stream of sound is strong relative to streams that are being ignored. Importantly, if attention is focused on a sound from a particular direction vs. focused on sound with particular non-spatial sound features, seemingly similar behavioral tasks actually engage very different brain networks. The network engaged by spatial auditory attention includes prefrontal and parietal areas, while non-spatial attention engages regions associated with high- level auditory processing. By contrasting fMRI activity during comparable auditory and visual selective attention tasks, we find that the cortical networks engaged by spatial auditory attention map directly to regions that are commonly assumed to comprise a visuo-spatial attention network. Conversely, the prefrontal regions that appear to be preferentially engaged ruing auditory processing are recruited during visual tasks that require analysis of stimulus timing. Together, these results support the view that auditory inputs are naturally processed differently from visual inputs, but that the brain has the capacity to recruit different brain networks to process the same inputs, based on task demands.

**Email:** shinn@cns.bu.edu

**11:00**

## The effects of action and attention on auditory learning

**Elizabeth Casserly*** *Trinity College*
**Hunter Drews** *Trinity College*

A major effort has been undertaken in recent years to develop better training methods for improving auditory perception in hearing-impaired populations and cochlear implant users. These individuals vary widely in their functional hearing capacity, and training protocols based on improvements in low-level discrimination ability have had limited success in producing real-world perceptual gains. We expand these efforts by testing the generalization of learning produced via two well-known mechanisms: self-produced action and engaged attention. Sixty college students with normal hearing were tested using simulation of cochlear implant acoustic processing. Participants completed a one hour training session in one of four conditions, with two pre/post-training speech recognition tests (in quiet/noise) and three additional generalization post-tests. In Experiment 1, two sets of control learners (sham/standard keyboard training methods) were compared to learners who produced target training materials aloud while receiving auditory feedback through real-time cochlear implant simulation, thereby completing the relevant action/perception loop. In Experiment 2, the same controls were compared to learners whose training materials (celebrity interviews) were designed to be relatively engaging in order to boost sustained attention. As predicted, both kinds of experimental learning resulted in significant post-training improvements. Specifically, Experiment 1 found gains in speech perception in noise, a notoriously difficult skill for those with hearing impairment, and in listeners' use of semantic context. Experiment 2 also showed benefits beyond the control in the use of context, as well as dealing with talker variability and recognizing speech in quiet. Moreover, a self-evaluation of attention showed that participants exposed to interviews were more engaged in the last third of training than any other group, and that attention self-ratings in all groups were significantly correlated with post-training scores on 4/5 perceptual tests. The application of classic perceptual learning effects to clinical auditory rehabilitation, therefore, appears to have promising potential.

Email: elizabeth.casserly@trincoll.edu, drews@geneseo.edu

**11:15**

## The effect of auditory cues on co-modulation masking release: The role of attention in the detection of sounds masked by narrow-band noise

**Doug Alards-Tomalin*** *University of Manitoba*
**Todd Mondor** *University of Manitoba*
**Jason Leboe-McGowan** *University of Manitoba*

Detecting a pure sound masked by amplitude fluctuated noise is improved by the presence of overlapped, spectrally distant, flanking sounds exhibiting the same envelope function as the mask, a phenomenon called co-modulation masking release (CMR). While most studies have tested situations under which CMR can, or cannot, be reliably elicited, fewer have examined the boundary limitations within the phenomenon itself. Currently, we examined the effect of preceding auditory cues on the degree of CMR. It has been suggested that the benefit of modulated flanking sounds (i.e., decrease in detection threshold), is largely informational in nature. For example, Information about the masking noise's envelope function is thought to be acquired from the spectrally flanking sounds, which allow people to be more effective at listening in the dips of the mask for the signal. In an earlier study, Bernstein and Trahiotis (1994) presented the flankers prior to the masked target. They found that improved signal detection occurred only when flankers exhibited frequencies similar to the signal. In our study, participants detected the presence of a 200 ms target signal that was either 600 or 900 Hz, embedded in a 1000 ms burst of narrow-bandwidth filtered Gaussian noise. While the signal had a constant amplitude, the masking noise was modulated to follow a sinusoid, and either contained, or did not contain flankers. In all of the experiments a 300 ms auditory cue was presented prior to the masked signal. In Experiment 1, the cues were either valid (same as the signal frequency) tones, or invalid (different from the signal frequency) tones. In Experiments 2 and 3, the cues contained information about the mask instead of the signal. While cueing the signal overall reduced the degree of CMR versus control conditions, valid cues that signaled information about the mask increased CMR over invalid cues.

Email: douglas.alards-tomalin@umanitoba.ca, todd_mondor@umanitoba.ca, jason.leboe-mcgowan@ad.umanitoba.ca

**11:30**

## The impact of working memory load and noise on the online processing of spoken words: Evidence from eye movements

**Boaz Ben-David*** *Interdisciplinary Center Herzl*
**gal Nitsan** *Interdisciplinary Center Herzl*
**Daniel Algom** *Tel-Aviv University*
**Arthur Wingfield** *Brandeis University*

In daily life, speech perception is usually accompanied by a noisy background and other tasks that tap into working memory capacity. However, the role of working memory (WM) on speech perception in noise is not clear. The goal of this study was to examine how WM load affects the timeline for spoken word recognition in noise. We used eye-tracking ('Visual World' paradigm) as an on-line measure of spoken word recognition. Listeners heard spoken instructions that relate to an object presented in the visual display, while their eye movements were recorded. For example, participants might be asked to "touch the candle," as the visual display contains four objects: candle, candy, dog and bicycle. As the speech signal unfolds, several alternatives are activated in response to phonemic information, i.e., CAND leads to candy and candle. In order to successfully achieve word identification, one has to inhibit phonological alternatives, once contradictory information is accumulated (DY). With millisecond accuracy, we recorded the extent to which the focus on the target, candle, is delayed due to competing activation of the onset competitor, candy, as reflected by the listener's gaze pattern. Thus, eye- movements captured listeners' ability to differentiate the target noun from its depicted phonological competitor.
We manipulated WM load by using a digit pre-load task: Participants retained either one (low-load) or four (high-load) spoken digits for the duration of the spoken-word recognition trial. Our first study showed that, in quiet, the ability to discriminate between the two phonologically competing alternatives was harder in the high-load than in the low-load condition. In the second study, words were presented on the background of speech-spectrum-noise (SNR= - 4dB). Initial data show that WM load had a larger effect on word recognition when words were presented in noise than in quiet. Individual differences in WM capacity also affected word recognition.

Email: boaz.ben.david@idc.ac.il, galnits2@gmail.com, algomd@post.tau.ac.il, wingfiel@brandeis.edu

**11:45**

## Spatial release from informational masking modulates auditory evoked potentials

**Benjamin Zobel*** *University of Massachusetts, Amherst*
**Lisa Sanders** *University of Massachusetts, Amherst*
**Richard Freyman** *University of Massachusetts, Amherst*

EEG was recorded while listeners detected noise-vocoded words (targets) presented with noise-vocoded two-talker speech (maskers). A target and masker were always presented from a loudspeaker positioned directly in front of the listener. Sometimes an identical copy of the masker was presented from another loudspeaker positioned 60 degrees to the listener's right with the onset of the right masker preceding the onset of the front masker by 4 ms. Although this latter arrangement adds masking sound, it produces the precedence effect for the masker and the perception of spatial separation between the target and masker. Consistent with previous behavioral research, target detection threshold was reduced by around 20 dB with the spatial cue of the precedence effect, indicating a substantial release from informational masking. For the first time, ERPs were used to measure differences in the auditory processing of targets across these conditions. When the masker was presented from both loudspeakers, targets unmasked by the precedence effect elicited clear auditory evoked potentials. However, when the masker was only presented from the front loudspeaker, these same targets largely failed to elicit the so- called obligatory auditory ERP response. These novel results suggest that the perception of spatial separation between targets and maskers dramatically facilitates the representation of target sounds at early stages of auditory processing. Implications for the role of attention in the spatial release from informational masking will be discussed.

Email: bzobel@psych.umass.edu, lsanders@psych.umass.edu, rlf@comdis.umass.edu

***Poster Session (12:00 noon – 1:20 PM)***
***Ballroom A – C (located on Level 3)***
***Abstracts located on pages 21 – 35***

***Lunch (1:20 – 2:00 PM)***

**2:00**

## *Invited Address*

## *A Revised Duplex Model for Sound Localization*

**William M. Hartmann** *Michigan State University*

Sounds are localized in the horizontal (azimuthal) plane by human (and other) listeners by means of interaural time differences (ITD) and interaural level differences (ILD). The best known model describing how these two interaural differences combine to create a localized image is the Duplex Model, which essentially says that at low frequencies the ITD is most strongly weighted and at high frequencies the ILD is most strongly weighted. The model is most readily applied to pure tones where the concept of frequency is unambiguous. Applications of the Model then predict localization as a linear combination of ITD and ILD with frequency-weighted coefficients. Recently, my colleagues and I have studied the localization of pure tones in free field (anechoic room) using transaural synthesis whereby tones with arbitrarily chosen parameters can be virtually created in the listener's ear canals without using headphones. Therefore, listeners localize sounds – even artificially created sounds. Experiments using artificial stimuli with natural ITDs and fixed ILDs that were opposite in sign to the ITDs revealed an entirely unexpected mode of binaural processing. Given an adequate opposing ILD, the binaural system reinterpreted the ITD information by seizing upon the ITD removed from the true ITD by a full period of the tone (slipped- cycle ITD) – in the direction favored by the ILD. Analysis of the data showed that the principle role of the opposing ILD was to switch between alternative interpretations of the ITD. Remarkably, experiments with opposing ILDs of 6 dB and 12 dB led to no significant difference in localization for the two ILD values. In the frequency regime where this perceptual mode applies (750 Hz and above) the combination of ITD and ILD became a highly nonlinear one with ILD serving more as a switch than as a linear contributor to localization.

**Email:** hartmann@pa.msu.edu

**2:30 – 3:15**

## Panel Discussion: Unifying principles and individual differences: What should we focus on to understand speech?

**Joseph Toscano*** *Villanova University*
**Bob McMurray** *University of Iowa*
**Emily Myers** *University of Connecticut*
**Matt Goldrick** *Northwestern University*

Over the past several decades, speech researchers have largely focused on a few key issues in attempting to describe how listeners recognize speech (e.g., lack of invariance, top-down effects, perceptual adaptation). Many models have been proposed to explain these phenomena, often arguing for a set of general principles that characterize speech perception. Articulatory-based models provide elegant solutions to the problem of coarticulation. Auditory-based models provide equally elegant explanations for why non- speech sounds influence phonological categorization. Bayesian approaches demonstrate that optimal statistical learning principles can characterize perceptual adaptation in speech.

Thus, on the one hand, we have a range of models that offer unifying principles for describing speech perception. At the same time, the field is filled with cases that seem to defy explanations based on a single set of perceptual principles. Why do some listeners appear to show categorical perception, while others demonstrate non-categorical perception, and still others show differences depending on the specific phonetic contrast? Why are some listeners readily able to adapt to novel accents, but others cannot? Why does lexically-guided perceptual learning work for some speech contrasts but not all?

Models must ultimately explain these idiosyncratic effects, not just the coarse set of canonical phenomena. Thus, perhaps we've been looking in the wrong places if our goal is to explain how listeners recognize speech. Could we develop better models by focusing on understanding idiosyncratic differences between listeners? Or, is the field still too immature to produce a more general model? If there are no unifying principles, what does this mean for the questions we ask about speech and auditory perception more generally?

Panelists will present work from speech perception and production that can provide insights into these questions and could lead to new ways of understanding individual differences and core principles that characterize how listeners recognize speech.

**Panelists:**

*Joe Toscano* will discuss work demonstrating idiosyncratic effects of specific speech sounds, in terms of the acoustic cues listeners' use and their ability to recognize speech in noise.

*Bob McMurray* will discuss work demonstrating that the way asynchronous cues are integrated seems to depend on the specific phonetic contrasts being studied, and work showing reliable individual differences in how categorically speech is perceived.

*Emily Myers* will discuss whether individual differences suggest that there are no unifying principles, or whether, despite individual variability, there is a common framework that can account for these differences.

*Matt Goldrick* will discuss how related work from speech production informs these issues, and whether or not the lack of unifying principles could be due to the immature state of cognitive science as a whole.

Email: joseph.toscano@villanova.edu; bob-mcmurray@uiowa.edu; emily.myers@uconn.edu; matt-goldrick@northwestern.edu

**3:15 Break (15 mins)**

**3:30**

## Looming sounds are perceived as faster than receding sounds

**John Neuhoff*** *The College of Wooster*
**Christian Golden** *The College of Wooster*

Moving motor vehicles injure and kill hundreds of thousands of pedestrians each year. This situation has the potential to get dramatically worse as both the production of quieter electric and hybrid vehicles and the number of pedestrians distracted by hand-held electronic devices increase. However, a better understanding of the perceptual and cognitive aspects of looming sounds could lead to a reduction in these fatalities. Throughout our evolutionary history looming objects have posed a threat to survival, and perceptual systems have evolved unique solutions to confront these environmental challenges. Vision provides an accurate representation of time-to-contact with a looming object and usually allows us to interact successfully with the object if required. However, audition functions as a warning system and yields an anticipatory representation of arrival time, indicating that the object has arrived when it is still some distance away. The bias provides a temporal margin of safety that allows more time to initiate defensive actions. In three studies this bias was shown to influence the perception of the speed of looming and receding sound sources. Listeners heard looming and receding sound sources and judged how fast they were moving. Listeners perceived the speed of looming sounds as faster than that of equivalent receding sounds. Listeners also showed better discrimination of the speed of looming sounds than receding sounds. Finally, close sounds were perceived as faster than distant sounds. The results suggest a prioritization of the perception of the speed of looming and receding sounds that mirrors the level of threat posed by moving objects in the environment.

Email: JNEUHOFF@wooster.edu, cgolden@wooster.edu

**3:45**

## Hearing gestures and drawing sounds: Auditory and multisensory perception of biological movements

**Etienne Thoret*** *LMA-CNRS, UPR7051, Aix-Marseille Univ, Ecole Centrale Marseille*
**Mitsuko Aramaki** *LMA-CNRS, UPR7051, Aix-Marseille Univ, Ecole Centrale Marseille*
**Sølvi Ystad** *LMA-CNRS, UPR7051, Aix-Marseille Univ, Ecole Centrale Marseille*
**Richard Kronland-Martinet** *LMA-CNRS, UPR7051, Aix-Marseille Univ, Ecole Centrale Marseille*

Understanding the links between the acoustical properties of sounds produced by our actions and the properties of the action imagined from them is a challenge that may help to understand how our sensorimotor interaction with the surrounding world constrains our perception and reciprocally. Here we'll present results of recent works on the acoustic characterization of a particular category of biological movements, namely drawing movements. If we listen carefully to the friction sounds produced by someone drawing, a movement can be perceived. How do we interpret this movement perceived from a friction sound? Is it for instance possible to recognize whether or not it is a human movement and can the drawn shapes be deduced from the dynamics of such sounds? To investigate these questions, we developed sound synthesis tools enabling the generation of virtual friction sounds which can be precisely calibrated according to movements parameters such as the velocity. A first experiment showed that we are able to recognized biological drawing movements characterized by the 2/3 power law through kinematic related friction sounds (Viviani et al., JEP:HPP 1992, Thoret et al., JEP:HPP 2014). In a second time, we showed that kinematic related friction sounds enable the recognition of geometrical shapes and even to characterize the prototype (Rosch, 1973) of an elliptical shape. Finally, in a multimodal context, we showed that friction sounds evoking movements may distort the visuo-motor coupling of drawing movements (Thoret et al., PLoS one, 2016). Taken together, all these experiments reveal the role of the auditory perception in the perception and production of biological movements. Moreover, it provides a new framework to investigate the relationship between sounds and movements.

Email: etienne.thoret@mcgill.ca, aramaki@lma.cnrs-mrs.fr, ystad@lma.cnrs-mrs.fr, kronland@lma.cnrs-mrs.fr

**4:00**

## Music expertise alters the perception of auditory objects

**Mrinmayi Kulkarni*** *University of Wisconsin Milwaukee*
**Katherine Potkonjak** *University of Wisconsin Milwaukee*
**Richard Randall** *Carnegie Mellon University*
**Adam Greenberg** *University of Wisconsin Milwaukee*

Low-level properties of auditory stimuli can alter the perception of auditory objects. The present study aimed to understand if high-level cognitive abilities, such as the level of formal music training, affect auditory object perception. Forty participants were categorized based on their responses to a survey which assessed level of music training through a variety of measures (formal music education, sheet music reading ability, etc.). These individuals were then tested behaviorally using stimuli comprising 100 randomly generated pure tone sequences, restricted to a one octave, Major, diatonic scale. Following the presentation of each stimulus, participants rated perceived musicality using a five-point Likert scale (1 = not musical; 5 = very musical), where musicality ratings served as indirect indices for the level of "objectness" of the musical sequences. A Principal Components Analysis (PCA) applied to participants revealed significant components that correlated with the degree of music training. On the other hand, PCA applied to stimuli revealed components that correlated with stimulus music properties (mean interval, key, contour, etc). Additionally, comparisons between subject groups revealed variable patterns in the ratings of participants with different levels of music training. Whereas, participants with high levels of music training consistently used a subset of these music properties while rating the musical sequences, participants with minimal or no music training did not make use of any such consistent strategy. These results suggest that music expertise may cause changes in auditory attention (and associated neural mechanisms) by selecting which cues should be attended/ignored during auditory object perception.

Email: mrinmayi@uwm.edu, potkonj2@uwm.edu, randall@cmu.edu, agreenb@uwm.edu

**4:15**

## Comparing perception–action coupling in groups of improvising and non-improvising musicians with an EEG Stroop task

**Andrew Goldman*** *Columbia University*
**Paul Sajda** *Columbia University*

One way to effectively apply neuroscientific methods to the study of improvisation is to examine between-group differences between representations of musical structures. Despite having overlapping vocabularies of musical structures (e.g., chords and scales), jazz improvisers learn in different ways than do classical musicians, often emphasizing practices like imitation and transcription that link perception and action. Different ways of knowing about these structures, manifested through differences in perception and motor planning, can be probed with neuroscience. Using a novel Stroop task, we compared perception-action coupling in groups of jazz and classical pianists, categorized using a questionnaire. Participants played triads on a MIDI keyboard while EEG was recorded. The MIDI signal was manipulated according to three conditions: congruent, quality incongruent, and shape incongruent. There was also a listen-only condition. In all four conditions, participants identified the inversion of the chord they heard. Thus, a mismatch between the anticipated sound of the played chord and the actual sound of the heard chord (which sometimes was incongruent) was meant to elicit perceptual interference. Stronger anticipation means stronger interference. Separate ANOVAs comparing the congruent condition against the others revealed main effects of condition on the reaction time (greater for incongruence) and N450 ERP component (greater negativity for incongruence), suggesting this task effectively elicited a Stroop effect. Further, playing congruent chords facilitated perception compared to listen-only. Contrary to the hypothesis, the classical group showed more Stroop interference according to the N450, and more facilitation by playing versus listening only. One explanation for this is that for the improvisers, the incongruent chords may not have been as perceptually incongruent since the particular set of stimuli in this experiment had similar functional-harmonic properties. This observation suggests future studies (e.g., using mismatch negativity paradigms) that compare different ways different groups of musicians categorize musical structures.

Email: ajg2232@columbia.edu, ps629@columbia.edu

**4:30**

## Bass drop: Establishing a link between low-frequency tones and movement

**Michael Hove*** *Fitchburg State University*

Auditory and motor systems are intimately linked. Auditory rhythms can activate motor and timing networks and induce movement in listeners. Movement and timing are more influenced by some auditory stimuli than others. In a series of studies, we have established that movement and timing are especially influenced by low-frequency (bass) tones. Here, I will present an overview of our behavioral and neuroimaging studies that establish this link. In finger-tapping studies, we have shown that lower-frequency tones have a greater influence on tap timing (both in tap-to-tone asynchrony and error correction). We have shown that songs rated high on groove ('the musical quality that compels movement') had more spectral flux in bass frequencies. In a transcranial magnetic stimulation (TMS) study, musical stimuli with more bass increasingly affected the motor system (i.e., cortico-spinal excitability). Using EEG (mismatch negativity), we showed that the brain better detects timing deviations in lower-pitched tones, and modeling suggests that encoding in the auditory periphery contributes to this effect. A new study presented music over sound-isolating earphones with or without a subwoofer that produced tactile bodily sensations; results showed that the auditory+tactile condition yielded higher groove ratings and induced more movement. In sum, the link between low-frequency tones and movement that we've repeatedly observed has physiological underpinnings (in auditory encoding and somato-motor activation) and explains the musical convention for low-pitched instruments to lay down the rhythm.

Email: michaeljhove@gmail.com

**4:45**

## A mutual adaptation model of joint action in music performance

**Alexander Demos*** *University of Illinois at Chicago*
**Hamed Laeghi** *McGill University*
**Caroline Palmer** *McGill University*

For joint action to occur in music performance, each individual must both adapt to and anticipate the heard actions of their partner. Adaptation and anticipation have often been understood in terms of forward-models in which individuals predict the actions of their partner based on simulations of the feed-forward connections between the motor and sensory systems. We propose a different approach, conceptualizing joint actions as a nonlinearly coupled feed-back system. From this dynamical systems perceptive, we have developed a computational model in which each person is represented by a simple linear oscillator each coupled to the other (i.e., sharing information). Each oscillator also receives delayed feedback from its own actions. The coupling of the two feedback oscillators creates a de-facto non-linear system. The two features of this system i.e., the delayed self-feedback of each oscillator interacted with coupling between oscillators, results in the emergent behavior of each oscillator anticipating the other. The directionality (who leads and who follows) can be changed by adjusting the degree of coupling of each oscillator individually. The model simulations were compared with human data of pianists performing duets in a perturbation study in which sounded feedback was randomly removed/returned from the parts performed by one or both partners. Model simulations showed similar behavior to human behavior in the amount of anticipation, the change in directionality caused by the perturbations, and finally, the time to recover after sounded feedback was returned. In short, the anticipation and adaptation in music performance can be modeled as a simple feedback system without the need to rely on forward-models and cognitive prediction/simulation.

Email: ademos@uic.edu, hamed.layeghi@mail.mcgill.ca, caroline.palmer@mcgill.ca

**5:00** **Updates & Closing Remarks**

# Poster session (12 noon – 1:20 PM)
# Ballroom A – C (located on Level 3)

1

## Subcortical encoding of sound: Interaction between attentional demand and stimulus regularity

**Zilong Xie*** *University of Texas at Austin*
**Bharath Chandrasekaran** *University of Texas at Austin*

The encoding of sound in the subcortical auditory nuclei has shown to be modulated by top-down attentional demands and the regularities of bottom-up inputs. However, the influences of these factors on subcortical auditory encoding have rarely been considered together. Here we investigated the interaction between attentional demand and input regularity in modulating subcortical auditory encoding. Participants performed a visual search task varying in attentional demand. On a random subset (50%) of trials, task-irrelevant pitch patterns were presented concurrently with the visual stimuli under low or high predictability conditions. We assessed subcortical electrophysiological responses to the pitch patterns. Results revealed that higher visual attentional demand was associated with decreased subcortical encoding of pitch patterns in the high predictability conditions, but the trend reversed in the low predictability conditions. These findings highlight an interaction between top-down attentional control and stimulus regularity in determining the fidelity of subcortical encoding of sound.

Email: xzilong@gmail.com; bchandra@utexas.edu

2

## Developing an auditory and visual cross-modal continuous performance task for evaluating concussion

**Christopher Koch*** *George Fox University*
**Taylor Charbonnier** *George Fox University*
**Kristin Dissinger** *George Fox University*
**Steven Egeberg** *George Fox University*
**Matthew Johnson** *George Fox University*
**Lindsay Levanen** *George Fox University*
**Matthew Scott** *George Fox University*

Neurocognitive tests like the SCAT3 and ImPACT have become standard concussion assessment tools. Although these tests have adequate sensitivity, specificity, and reliability, they are unimodal in nature. Consequently, the tests do not fully assess the range of processing that can be affected by concussion (Thompson, 2012). Therefore, we developed a cross-modal continuous performance task to examine cognitive processing post-concussion. Forty-three middle school school lacrosse players, college students, and physical therapy graduate students participated in the study. Twelve of these participants had been previously diagnosed with a concussion. Participants completed a symptom checklist from SCAT3 along with other demographic information (e.g., previously concussed, last concussion). They then completed the continuous performance task starting with visual detection followed by visual inhibition, auditory detection, and auditory inhibition. Older subjects were more accurate than younger subjects on the detection task ($F(1, 84) = 20.61$, $p < .001$). Subjects were also more accurate on the visual task than the auditory task ($F(1, 84) = 21.47$, $p < .001$). Both age ($F(1, 84) = 5.65$, $p < .02$) and previous concussion ($F(1, 84) = 4.49$, $p < .04$) interacted with test modality. College and graduate students who had previously been concussed performed the same as those who had not been concussed. However, middle schoolers who had been concussed did significantly worse on the auditory task than those who had not been concussed. Similarly, older subjects were more accurate than younger subjects on the inhibition task ($F(1, 84) = 4.91$, $p < .03$). Older subjects were also significantly more accurate on the visual task than the middle schoolers ($F(1, 84) = 5.33$, $p < .03$). However, no differences were found based on previous concussion.

Email: ckoch@georgefox.edu; tcharbonnier14@georgefox.edu; kdissinger14@georgefox.edu; segeberg14@georgefox.edu; matthewj14@georgefox.edu; llevanen14@georgefox.edu; mscott14@georgefox.edu

3

## More attention, less imitation?

**Marie Postma*** *Tilburg University*

Using the standard STOP-IT response inhibition task and a dichotic listening task to measure attention shifting, we explored the relation between individual selective attention abilities and pitch imitation in speech. In our study, participants performed a vocal repetition task under memory load by iterating auditory number words presented to them by an Embodied Interaction Agent while simultaneously trying to remember a sequence of numbers. The vocal output in the repetition task was compared to the speaker's baseline for the same number words. In order to obtain a comprehensive measure of pitch imitation, we analyzed three acoustic measures (mean f0, SD f0, and f0 slope). In Experiment 1 (*N* = 83), we compared the vocal imitative performance of two age groups to their average stop-signal reaction time. We found an age- related effect on both the imitative performance and the response inhibition but no correlation between the two. The vocal output of older adults was more likely to converge towards the f0 characteristics of the model talker, while their inhibition performance was poorer compared to the performance of younger adults. In Experiment 2 (*N* = 75), we used a variant of the auditory shifting task with dichotic listening to estimate participant's cognitive flexibility with respect to auditory stimuli in three groups of participants with different professional experience. There was a negative correlation between attention shifting and vocal imitative performance; furthermore, professional experience had an effect on attention shifting abilities. Together, our findings suggest that vocal imitation is related to some attention control abilities (i.e., shifting but not inhibition).

Email: Marie.Postma@uvt.nl

4

## Effects of ruminative mind wandering episodes on speech imitation, lexical associations and eye behavior

**Marie Postma*** *Tilburg University*

In two studies, we investigated the effect of repetitive, self-focused ruminative episodes on eye behavior and lexical associations (Study 1) and eye movement and vocal imitative behavior (Study 2) in a controlled game-like environment approximating human interactions. Participants (*N* = 61) were randomly assigned to a rumination-induction or a control condition; they played a word association game with a male Embodied Interaction Agent during which their eye movements and verbal and non-verbal behavior were recorded. Two questionnaires were used to measure their ruminative traits and task-unrelated thoughts. Relative to the control condition, participants who were induced to ruminate tended to mind wander more and had more negative lexical associations triggered by semantically neutral lexical items. With respect to eye movements, rumination was associated with more eye blinks. Interestingly, participants in the rumination condition also displayed less vocal imitative behavior, suggesting that the inward focus of ruminative mind wandering may have an impact on automatic processes accompanying social interactions. In sum, this research demonstrates that ruminative self-focus is reflected in two nonverbal signals that both play an important role in face-to-face interactions.

Email: Marie.Postma@uvt.nl

5

## The influence of musical instrumentation on the occurrence of audiovisual associative links

**Elena Anisimova*** *University of Music Karlsruhe*

A test was designed to detect the phenomenon of "arbitrary synesthesia" (cross-modal associative links that occur during the perception of music) for various audio-visual combinations of people who are not synesthetes. This test investigates not only the formation of audio-visual associative links, but also the connections between auditory and visual perception on a cognitive-emotional level. Additionally, it examines the influence of instrumentation on the occurrence of audiovisual associative links. A total of 110 individuals without synesthesia completed the test. Participants listened to musical excerpts that were each about 20 seconds in length, and then indicated any potential associations using descriptive terms. The obtained data indicated a correlation between sound perception and associated visual object (i.e., “arbitrary synesthesia”) that is based upon cognitive factors. Specifically, characteristics of musical timbre and their intensity were affected by audio-visual associative links. Musical instruments with a “rich” timbre were associated with “dark” colors, whereas instruments with a “clear” timbre were associated with “cool” and “light” colors, and instruments with a more distinct sound were associated with colors that reflect high values of lightness. Implications of these findings for understanding the relevant psychological dimensions in music that contribute to audio-visual correlations and their intensity will be discussed.

Email: MusicproductionStarpegasus@yandex.com

6

## The effect of stimulus frequency, quality, duration, and location on temporal order judgment threshold

**Leah Fostick*** *Ariel University*
**Harvey Babkoff** *Bar-Ilan University*

Auditory temporal order judgment (TOJ) is a widely used test for measuring auditory temporal processing. In this study, we examined the influence of stimulus parameters, e.g., frequency, quality, duration, and tone location, on TOJ threshold. Participants (*n*=192) were divided into seven groups. Three groups were presented with two pure tones (PT) of different frequencies (1 and 2kHz; 1 and 1.1kHz; 1 and 3.5kHz), a fourth group was presented with a 1kHz PT and a Gaussian noise (GN) burst. The fifth and sixth group were presented with two 1kHz tones of different duration (10 and 30msec; 15 and 45msec). The seventh group of participants judged the temporal order of two 1kHz tones presented asynchronously to the right and left ears (perceived as different locations). Very short TOJ thresholds of ISI=0msec (no separation between the offset of the first stimulus and the onset of the second stimulus) were found for 32-48% of the first three groups of participants who performed TOJ with different frequencies. Twelve percent of the participants who received a tone and a Guassian noise had very short thresholds. However, none of the participants who received tones of different duration or who received the same tone asynchronously to the two ears (location) had very short thresholds. All of the participants in the latter group (location) had TOJ thresholds of 0 < ISI < 250msec. For the other six groups, 20-40% of participants had thresholds > 250msec. These results suggest that the parameters that identify the two stimuli whose temporal order is to be judged, are significant factors in determining TOJ threshold, especially the ability to correctly judge TOJ when ISI=0 msec. Differences in stimulus parameters might provide cues other than the temporal information inherent in ISI, which can be successfully used by some of the participants to judge order.

Email: leah.fostick@ariel.ac.il; babkofh@gmail.com

7

## Phonetic category activation can drive rapid speech adaptation

**Yunan Wu*** *Carnegie Mellon University*
**Lori Holt** *Carnegie Mellon University*

Listeners are sensitive to correlations among the acoustic dimensions that define speech categories. Highly diagnostic dimensions have the greatest impact on categorization, but correlated secondary dimensions also contribute. Prior research has demonstrated that this relative 'perceptual weight' is very sensitive to regularities in short-term speech experience. When listeners encounter a correlation between dimensions that runs counter to long-term experience they rapidly down-weight reliance on the secondary dimension in speech categorization. Here, we test the hypothesis that phonetic-category- level activation via a highly diagnostic dimension is critical in driving this rapid tuning of how input maps to speech categories. We used noise-vocoding, a signal-processing technique, to manipulate the inherent relative informativeness of two acoustic dimensions significant to English vowel categorization. The pattern of results was driven by the effectiveness of an acoustic dimension in signaling vowel category membership, consistent with our hypothesis. These findings bear on models of speech categorization and rapid adaptive plasticity in perception, more generally.

Email: charleswu.01@gmail.com; loriholt@cmu.edu

8

## Happy mood makes music sound brighter

**Ryan Brigante*** *University of Texas at Dallas*
**Bart Rypma** *University of Texas at Dallas*

Grounded theories of cognition suggest that bodily states, including internal affective states, influence the way we perceive and understand the world. Conceptual metaphors such as "happiness is bright" and "sadness is dark" link different domains of experience, and the linguistic expressions are thought to reflect underlying mental mappings. In addition to visual-affective metaphors, we also have visual-musical metaphors such as "dark tone" and "bright harmony." Previous experiments found that perception of bright visual stimuli corresponded with brighter perception of musical timbre and harmony. Do affective states also alter our perception of music in a metaphorical fashion? If so, then feeling happier or sadder should correspond with perceiving musical stimuli as brighter or darker. We employed a mood induction procedure (MIP) to encourage participants to get into a happy or sad mood, and we tested their subjective judgments of timbre and harmonic brightness both before and after mood induction. The MIP involved listening to happy or sad music while reading uplifting or depressing sentences. The perceptual tests involved listening to individual tones or chords played by various orchestral instruments and judging their timbral or harmonic brightness on a subjective rating scale. The MIP had the intended effect, causing short-lasting alterations of happy and depressed mood in the expected directions (i.e., happiness increased and depression decreased following positive mood induction, and the opposite for negative mood induction). Participants in a positive mood judged the musical stimuli to be brighter than those in a negative mood. Crucially, there was no difference in baseline brightness judgments prior to mood induction when comparing the same participants. These results suggest that positive mood encourages a brighter perception of musical timbre and harmony, while negative mood encourages a darker perception of timbre and harmony. Thus, musical- affective metaphors seem to be reflected in perceptual judgments.

Email: ryan.brigante@gmail.com, bart.rypma@utdallas.edu

9

## Listeners anticipate upcoming vowels before identifying the current fricative.

**Kayleen Schreiber*** *University of Iowa*
**Bob McMurray** *University of Iowa*

Speech unfolds quickly over time. Consequently, not all of the information needed to identify a given phoneme will be available at any moment in time. Listeners must use some sort of strategy to integrate information across time to successfully perceive speech. Consider two possible strategies. Under a late integration strategy, initial cues are stored in a memory buffer, until all of necessary information has been acquired. At this point lexical items can be activated. In contrast, under an immediate utilization strategy, auditory information is used to begin activating possible words as soon as any information is available. Lexical activation is then updated continuously as later information arrives. A multitude of studies have examined this problem of integrating asynchronous cues (McMurray, Clayards, Tanenhaus, & Aslin, 2008; Miller & Dexter, 1988; Reinisch & Sjerps, 2013; Toscano & McMurray, 2012, 2015), and shown that listeners start using any speech information to activate lexical items as soon as they arrive, supporting an immediate utilization account. However, these studies have focused on stop consonants and vowels. More recently, our lab found that for words with initial fricatives, listeners adopt a late integration strategy (Galle, Klein, Schreiber, & McMurray, submitted). This unexpected result was replicated multiple times and seems to indicate that listeners use a different strategy to integrate cues for fricatives than other speech sounds. One possibility is that this derives from the auditory properties of fricatives, which are aperiodic and appear on very distinct frequency bands from vocalic portions of speech.

Whereas the prior studies focused on accumulating information to make a decision (retrospectively) about prior sounds, the present study asked how listeners integrate cues to make a prospective or anticipatory decision about upcoming sounds. In particular, coarticulatory information present in the frication can provide information about the upcoming vowel, specifically whether it is rounded or not (/i/ vs. /u/). For example, the /s/ before seed is pronounced differently than the /s/ before soup. We asked what strategy listeners employ to use the coarticulation in the frication to anticipate the upcoming vowel. To measure the temporal unfolding of lexical activation, we used eye-tracking in the Visual World Paradigm. Participants heard a spoken word and selected the corresponding picture from a screen containing pictures of the target (soup), a competitor (seed) and unrelated items (bug, pack). We manipulated the coarticulation in the frication, to create stimuli in which the fricative anticipated either a rounded or unrounded vowel. By measuring how strongly biased participant's eye movements were to the /i/ or /u/ over time, we estimated the unfolding anticipation and commitment to the vowel. Using this technique, we estimated when participants used the coarticulatory cues in the frication, and when they used the actual identity of the vowel (e.g., the /i/ or /u/) to influence lexical access. Listeners began using the coarticulation soon as it was available, and prior to the use of the vowel identity. This suggests an immediate utilization strategy for anticipation. In contrast, a second set of stimuli examined when people use those same frication portions to identify the fricative (e.g., soup vs. shoot). This replicated the result that listeners wait until the vowel to begin using the frication. Oddly then, listeners appear to anticipate the upcoming vowel before they have identified the current fricative. This result suggests that a late integration strategy does not derive from something inherent in the acoustics of frication (e.g., an auditory property). It also suggests that listeners employ different cue integration strategies for retrospective and prospective processing.

Email: kayleen-hannaway@uiowa.edu; bob-mcmurray@uiowa.edu

10

## Modeling audiovisual cue integration in speech

**Laura Getz*** *Villanova University*
**Elke Nordeen** *Villanova University*
**Sarah Vrabic** *Villanova University*
**Joseph Toscano** *Villanova University*

Adult speech perception is generally enhanced when information is provided from multiple modalities. In contrast, infants do not appear to benefit from combining auditory and visual speech information early in development. This is true despite the fact that both modalities are important to speech comprehension even at early stages of language acquisition. How then do listeners learn how to process auditory and visual information as part of a unified signal? In the auditory domain, statistical learning processes provide an excellent mechanism for acquiring phonological categories. Is this also true for the more complex problem of acquiring audio-visual correspondences, which require the learner to integrate information from multiple modalities? We asked whether simple statistical learning approaches are sufficient for learning such multi-modal representations. We simulated the developmental process of acquiring phonological categories from auditory cues (i.e., F2 and F3) and visual cues (i.e., mouth height and mouth width) using Gaussian mixture models that learn cue weights and combine cues on the basis of their distributional statistics. We find that the model can easily detect the intended phoneme category when auditory and visual inputs provide redundant information. From a developmental perspective, we find that combined audiovisual representations form later in development than cue-level representations. Additionally, early in development, auditory cues are used almost exclusively in cases where auditory and visual inputs are mismatched. Later in development, visual cues are given higher weights, leading to more ambiguous phoneme detection (as in the McGurk effect). Overall, domain-general statistical learning techniques allow us to model the developmental trajectory of audio-visual cue integration in speech, and in turn, allow us to better understand the mechanisms that give rise to unified percepts based on multiple cues.

Email: laura.getz@villanova.edu; enordeen@villanova.edu; svrabic@villanova.edu; joseph.toscano@villanova.edu

11

## Individual differences in dynamic adaptation during lexically-guided perceptual learning

**David Saltzman*** *University of Connecticut*
**Emily Myers** *University of Connecticut*

Across languages, it is necessary for listeners to compensate for variability in the speech signal. Lexically guided perceptual learning (LGPL) serves as a possible mechanism for listeners to adjust phonetic category criteria when confronted with non-standard phonetic exemplars. In LGPL, listeners shift their perception of a phonetic contrast in response to hearing an ambiguous token embedded in an unambiguous lexical context. Of interest is the degree to which individuals will re-shift their phonetic boundaries when confronted with conflicting information. Previous work (Kraljic & Samuel, 2005) shows that exposure to "good" versions of the ambiguous phoneme will re-set category, but these results are attributed to selective adaptation, and in general the durability of perceptual learning effects has been strongly highlighted. Nonetheless, a recent study has shown that listeners can continuously readjust category boundaries within an experiment in response to audio-visual as well as lexically biasing information (Ullas et al., 2016). The present study asked whether, using only lexically-biasing information, participants would readily shift their category boundaries within a session. Of further interest was the stability of these effects over time, as well as the consistency of individual differences in participants' readiness to adjust to new information. In the current study, participants were presented with a blocked design in which two blocks of a lexical decision task biased them in opposite directions of the /s/-/ʃ / distinction. Following each lexical decision task, participants were asked to categorize sounds along a continuum from "sign" to "shine". This task was then repeated between four to ten days later. Preliminary results show a robust main effect of /s/-/ʃ / bias, indicating a boundary shift in response to ambiguous tokens in the lexical decision phase. Follow-up analyses address the durability of such effects, and the degree to which individual differences in phonetic plasticity are consistent between sessions.

Email: david.saltzman@uconn.edu; emily.myers@uconn.edu

12

## Working memory capacity predicts the pupil dilation response to auditory deviation

**Alexandre Marois*** *Université Laval*
**Johnathan Crépeau** *Université Laval*
**François Vachon** *Université Laval*

Measures of working memory capacity (WMC) are frequently used as predictors of performance in a wide range of domains (e.g., Unsworth & Engle, 2007). For instance, high-WMC individuals appear more efficient at resisting the distractive power of a sound deviating from the auditory background (e.g., Hughes, Hurlstone, Marsh, Vachon, & Jones, 2013). While this relationship between WMC and the susceptibility to auditory attentional capture has been established at the behavioral level, it has never been tested using physiological indices of the orienting response (OR) to the deviant sound. The present study sought to determine whether individual differences in WMC can predict the size of the physiological response to auditory deviation. To do so, we assessed the variations in pupil size, known to be sensitive to deviant sounds (e.g., Steiner & Barry, 2011), in the context of the irrelevant sound paradigm whereby a deviant item was occasionally embedded within a to-be-ignored auditory sequence composed of the repetition of a standard sound while participants performed a visual serial recall task. For each of the 28 participants, the pupil dilation response (PDR) evoked by standard and deviant sounds was assessed and WMC was estimated using a complex span task (OSPAN). Results showed a significant negative correlation ($r$ = -.419, $p$ = .026) between the OSPAN score and the size of the deviation effect, calculated as the mean difference between the PDRs elicited by the deviant sounds and those evoked by the standard stimuli. The fact that higher WMC was associated with smaller physiological response to OR-triggering sounds is consistent with the idea that high WMC, reflective of better cognitive control, can attenuate the power of sounds to capture attention (Conway, Cowan, & Bunting, 2001)

Email: alexandre.marois.1@ulaval.ca; johnathan.crepeau.1@ulaval.ca; francois.vachon@psy.ulaval.ca

13

## Using sound to detect the presence of an occluding object: Can physics alone explain human perception?

**Stephanie Brown** *Washburn University*
**Michael K. Russell** *Washburn University*

Occlusion refers to those situations where one object partially or completely conceals another object. While occlusion has been repeatedly investigated in vision research, almost no research exists on the auditory perception of occlusion. This is particularly noteworthy given that, in everyday settings, sound-producing objects are often occluded in part or in whole. The beauty of sound (in opposition to light) is its ability to travel around objects, which allows us to hear sounds that are occluded. Previous research has revealed that sighted but blindfolded individuals are highly capable of detecting occlusion (partial or whole) and creating a state of occlusion (i.e., moving an object so that it occludes a loudspeaker in part or entirely). Physics suggest 3 possible explanations for our ability to perform such tasks: insertion loss, Fresnel number, and transmission loss. Insertion loss refers to the decrease in sound intensity at the point of observation that results from an object being placed between the observer and sound source. Of importance is the material composition of the occluding object. Fresnel numbers center on the idea that sound has the ability to travel around barriers and is a function of sound frequency and the distance between observer and sound source. Transmission loss, like insertion loss, refers to the decrease in sound energy and is based on the relationship between the material composition of the obstruct and the frequency of the sound signal. The present study examined the degree to which each of auditory occlusion perception (detection) is related to signal intensity, observer-source distance, and signal frequency. The results are discussed in terms of the apparent usefulness of insertion loss, Fresnel numbers, and transmission loss to be capable (and incapable) of explaining how it is that we are able to detect occlusion when relying solely on sound.

Email: mike.russell@washburn.edu; stephanie.brown4@washburn.edu

14

## Using echolocation to detect the existence of openings and object location

**Stephanie Brown** *Washburn University*
**Michael K. Russell** *Washburn University*

As is well known, echolocation refers to the emission of sound by a perceiving- acting organism and the use of reflected sound to judge the presence and position of objects, in particular, and spatial layout, in general. The ability to echolocate is particularly useful when one considers the need to be able to navigate through a cluttered world which often contains a relatively high proportion of objects that do not emit sounds. Generally speaking, humans are able to successfully echolocate such objects. However, research involving echolocation has tended to focus on the ability of individuals to detect the presence of objects. To echolocate successfully so that one may navigate in the real-word, individuals must be able to detect both objects and openings. We avoid objects and use openings to perform desired actions (e.g., walking). The ability to perceive openings is considered essential to perceiving-acting organisms since, according to James J. Gibson, openings permit opportunities for action. To date, no known study has investigated the ability of individuals to use echoes to detect openings. Participants in the present study were instructed to walk along a wall until they were standing in the center of the gap in that wall or until they stood adjacent to the sole closed portion of that wall. Participants were exposed to six target positions. The results revealed sighted-but-blindfolded individuals were largely unable to detect the gap location but were fairly accurate at judging object location. Gap location was affected by the sound signal but unaffected by experience. Object location ability was tied partially to experience. Discussion is given to the manner with which the acoustic structure is informative about an opening and how that acoustic information enhances our understanding of spatial layout.

Email: mike.russell@washburn.edu; stephanie.brown4@washburn.edu

15

## Examining the optimal and observed bounds of sound class discrimination

**Mattson Ogg*** *University of Maryland*
**Robert Slevc** *University of Maryland*

Auditory object identification relies on information that develops over time, yet nonetheless we identify sounds quickly, on the basis of only very limited acoustic information. It is not known what acoustic features listeners rely on to perform these tasks on such rapid time scales. It is also unclear what acoustic information is even available within the stimuli that can reliably discriminate classes of sounds at such short durations. Understanding this latter issue will provide context and an upper bound on object discrimination abilities early in auditory processing. Additionally, such information can provide new insights into how comprehensively previously identified acoustic features can characterize sounds in the context of such temporal constraints. Thus, we examined auditory object classification via machine learning classifiers which were trained and tested on category labels (speech, instrument, and human environment sounds) and a variety of acoustic features that have been effectively used in previous sound analysis and classification work (harmonic to noise ratio, spectral centroid, spectral flatness, spectral variability, acoustic power, etc.). Training and testing was performed via iterative cross-validation at durations starting from 12.5 *ms*, up to 200 *ms*. Instrumental stimuli were very reliably distinguished from speech and environmental sounds regardless of the time point tested. The classifier also performed above chance when discriminating speech and environmental stimuli, but did not perform as well as for instrument sounds, likely due to the larger acoustic variability within these signals. This observation was supported by further analysis of the weights the classifiers assigned to the acoustic features they were given. These feature weights will be compared to models of human performance on analogous tasks to examine whether human listeners are using the acoustic information available to them in an optimal manner.

Email: mogg@umd.edu; slevc@umd.edu

16

## Millenials write less-positive things than their parents about recent pop songs

**Anne Gilman*** *Juniata College*
**Clayton Reichart** *Juniata College*

Popular songs have been used to compare auditory imagery and perception (Halpern & Zatorre, 1999) and to delineate important life stages (Krumhansl & Zupnick, 2013). In order to predict how song stimuli might shape crossmodal working memory performance, 500ms song and other recordings were compared using criteria taken from earlier working memory and auditory cognition research. Response bias differences in auditory-visual change detection have been found for more-meaningful images, which evoked longer and more-varied verbal associations (Gilman, Ware, & Limber, 2010). In an effort to discover comparable effects for more- and less-meaningful sounds, 77 participants in two age cohorts (41 from 18 to 27 years old and 36 parents over 37) typed out what associations came to mind when listening to four types of sounds. Sixteen songs each were chosen from Billboard 1980's and 2000's charts, along with 16 ringtones, and 8 frequency-modulated chirp sequences from Sanders & Poeppel (2007). Length and variety of responses was comparable across all categories (Le et al., 2016). Could mining these associations provide emotion data comparable to ratings used in Krumhansl & Zupnick's study? Crowdsourced emotion ratings for participants' words (Warriner, Kuperman, & Brysbaert, 2013) were more positive for the songs than for the ringtones or the chirps, ($F(3,5981)=54.83$, $p<.001$). Interestingly, the older cohort's words about millenial songs (released 2000-2009) were more positive on average than the words provided by younger adults ($F(1,5981)=17.82$, $p<.01$), even though the older group did not recognize as many of the millenial songs (Causer et al., 2016). Concreteness ratings (Brysbaert, Warriner, & Kuperman, 2014) were used to relate to Krumhansl & Zupnick's (2013) measures of personally-relevant memories. Eighties music and chirps elicited the most concrete words ($M=3.34,3.30$); millenial tunes and ringtones ($M=3.20,3.21$) were significantly lower, ($F(3,9799)=7.971$, $p<.001$). Refinements of these measures could facilitate investigations of songs' working memory impact.

Email: anne.gilman@gmail.com; reichca13@juniata.edu

17

## Selective and divided attention in polyphonic melodies: Training related advantages

**George Seror III*** *University at Albany*

Polyphonic melodies were compared in a same-different task. In four counterbalanced experiments, effects of consonance (changes resulting in consonant vs. dissonant harmonic intervals), tonality (in-key vs. out-of-key changes), and voice (changes occurring in the upper vs. lower voice) were examined. The magnitude of change for consonance vs. dissonance (one-semitone vs. two-semitones) and direction of change from standard to deviant melodies (ascending vs. descending) were counterbalanced between experiments. Overall effects of consonance, tonality, and voice were found, when controlling for magnitude and direction of change. In addition, the magnitude of these effects covaried with years of music experience. The results are discussed in terms of training related advantages in selective and divided attention.

Email: gseror@albany.edu

18

## The influence of inversions on the perception of major and minor chords.

**Vincenzo Belli*** *University at Albany, SUNY*
**Abigail Kleinsmith** *University at Albany, SUNY*
**W. Neill** *University at Albany, SUNY*

A musical chord consists of three or more simultaneous notes. Of particular importance in Western music are “major” and “minor” triads. A root-position major triad consists of a tonic note, a note a major third (4 half-steps) above it, and a note a perfect fifth (7 half- steps) above it—e.g., C-E-G for a C major chord. Similarly, a root-position minor triad consists of a tonic, a minor third (3 half-steps) above it, and the perfect fifth—e.g., C-Eb- G for a C minor chord. Even non-musicians reliably distinguish between root-position major and minor triads as “happy” or “sad” respectively. However, there has been little investigation of how perception of major versus minor is affected by inversions: In a first inversion, the tonic is raised above the third and fifth (e.g., E-G-C); in a second inversion, both the tonic and third are raised above the fifth (e.g., G-C-E). In Experiment 1, subjects learned to judge triads as “major” or “minor”, with feedback on each trial. They discriminated second-inversion triads as well as root-position triads, but did poorly on first-inversion triads. In Experiment 2, subjects simply judged major or minor triads as sounding “happy” or “sad”, without feedback. The results mirrored the first experiment: Major triads were judged most often as “happy” and minor triads most often as “sad”, in root position and second inversion. However, in first inversion, more major triads sounded “sad” and more minor triads sounded “happy”. The results imply that perception of major and minor chords is not determined merely by pitch classes, but instead depends on the specific intervals between notes.

Email: vbellialbany@gmail.com; akleinsmith@albany.edu; wneill@albany.edu

19

## Which cues do listeners use? Discovering networks of phonetic cues for speech sound categorization using a graph theoretic approach

**Anne Marie Crinnion*** *Harvard University*
**Beth Malmskog** *Villanova University*
**Joseph Toscano** *Villanova University*

Identifying phonetic cues used by listeners is central to developing viable models of speech perception. However, current models require researchers to make assumptions about what those cues are. We present a novel solution to this problem based on concepts from graph theory. Specifically, we apply Steiner tree algorithms, recently used to study gene-protein networks, with the goal of extracting relevant acoustic cues for speech sound categorization. The model uses acoustic measurements (N=23 cues) and fricative classification data (*N*=20 listeners) from McMurray and Jongman (2010, Psychol Rev). Cue-values were computed by dividing cues into ten bins spanning the range of acoustic values across the eight fricatives (/f,v,,ð,s,z,ʃ,ʒ /) and assigning a code to each cue-bin combination, creating 230 unique codes. Next, graphs were created for each fricative with nodes corresponding to the 20 listeners and the 230 codes, with edges weighted by the inverse likelihood of indicating that fricative given the code (e.g., if a listener identified 50% of tokens with a given cue-value as an /s/, a weight-2 edge would connect the listener and that cue-value in that graph). Algorithms implemented in SageMath were used to find Steiner trees, which are subgraphs that minimize edge weights while connecting all listeners. Codes appearing in the tree were then eliminated, and the algorithm was run again. This process was repeated until no subgraph connecting all listeners could be found. Across the dataset, 107 cue-values were identified as connecting all listeners, suggesting that a large number of cues provide information for recognizing specific sounds in the dataset. Moreover, examination of the relevant cues reveals that most uniquely identify just one or two fricatives, indicating that they are highly informative. These data offer a new approach for identifying phonetic cues and suggest that massive cue integration principles may lead to new models of speech recognition.

Email: acrinnion@college.harvard.edu; beth.malmskog@villanova.edu; joseph.toscano@villanova.edu

20

## Measuring working memory: A theoretical exploration of musical working memory

**Juan Ventura*** *Louisiana State University*
**Emily Elliott** *Louisiana State University*

Working memory (WM) has been extensively studied, in part because of its strong relationship with higher-order cognitive abilities such as fluid intelligence (e.g., Kane et al., 2004; Shelton, Elliott, Matthews, Hill, & Gouvier, 2010). A core debate is modeling WM either as a unitary, domain-general system (e.g., embedded process model; Cowan, 2005) or a non-unitary system with domain-specific stores (e.g., multicomponent model; Baddeley, 1986). The verbal and visual domains have historically been the central domains of interest. By contrast, considerably less attention has been placed on the musical domain. Although the discussion of a musical WM system is not new, the issue has been largely outside of the mainstream focus, and in result, has been less developed than its visual and verbal counterparts. A musical WM system has been proposed with distinctive features when compared to a verbal system (Berz, 1995) and may potentially only exist in musicians (Schulze et al., 2011). In addition, there is growing research suggesting musical training enhances non-musical cognitive processes (e.g., Kraus & Chandrasekaran, 2010), especially in childhood. Within the literature, numerous approaches have been taken to investigate musical WM. The methodology and difficulty of task materials, participant variability, criteria of non-musician and musician, group vs individual difference analyses, and definitions of WM and short-term memory all vary from study to study. Consequently, these differences create difficulties when comparing findings across studies and when making advancements in answering questions about the development, structure, and enhancement of WM. This theoretical presentation will discuss past, present, and future approaches in examining and modeling musical WM, and will propose a measure that limits long-term memory strategies from influencing performance, limits sensory memory effects, and is analogous with common verbal and visual measures.

Email: jventu4@lsu.edu; eelliott@lsu.edu

21

## The interaction of repetition and short-term memory in melodic dictation tasks

**Nathan Cornelius*** *Peabody Institute*
**Jenine Brown** *Peabody Institute*

Several earlier studies have examined the effect of repetition in melodic dictation exercises in the undergraduate ear-training classroom, but findings have been contradictory. Some studies have found no significant correlation between the number of times participants heard a melody and their ability to notate it correctly (Hofstetter, 1981; Langsford, 1959; Sloboda and Parker, 1985); others have suggested that hearing a melody more times increases dictation accuracy levels (Pembrook, 1986). Our research builds upon these studies by specifically investigating the effect of repetition in lengthier melodic dictation tasks to more closely simulate the testing conditions of a typical dictation exercise in an undergraduate ear-training class. Untrained freshman music majors heard paired melodies of equal difficulty either three or six times and were asked to notate the melody. Mean performance was significantly higher when participants heard melodies repeated six times than three, offering some clarification to the empirical research previously cited. Several findings also raise important pedagogical implications. First, we found evidence that the variance in musical memory among untrained freshmen is wider than Karpinski (2000) hypothesizes, challenging the ear-training instructor to offer exercises at levels appropriate for the entire class. Second, this study offers the first empirical evidence of the dictation ability of incoming undergraduate freshmen music majors, inviting a long- term study on the extent to which musical memory increases during the four-semester ear-training curriculum. Third, while more repetitions of the melody improved scores for pitch, rhythm scores remained accurate throughout, suggesting that rhythmic complexity in ear-training exercises could be introduced sooner relative to the sequence found in seminal ear-training texts (e.g., Rogers and Ottman, 2013). Finally, dictation scores strongly correlated with years of piano instruction and less strongly with years of instrumental instruction, reinforcing the importance of incorporating keyboard skills within the ear-training curriculum.

Email: nathan.cornelius1@gmail.com; jenine.brown@jhu.edu

22

## Influence of spectrotemporal modulations on the recognition of musical instrument sounds

**Etienne Thoret*** *Schulich School of Music, McGill University, CIRMMT*
**Philippe Depalle** *Schulich School of Music, McGill University, CIRMMT*
**Stephen McAdams** *Schulich School of Music, McGill University, CIRMMT*

Modulation Power Spectra have been shown to contribute significantly to the perception of musical instrument sounds. Nevertheless, it remains unknown whether each instrument's identity is characterized by specific regions in this representation. Two recognition tasks were applied to two subsets of musical instruments: tuba, trombone, cello, saxophone, and clarinet on the one hand, and marimba, vibraphone, guitar, harp, and viola pizzicato on the other. The sounds were processed with filtered spectrotemporal modulations. Based on a "molecular approach", the so-called bubbles method, the most relevant parts of this representation for instrument identification were determined for each instrument and reveal the regions essential for their identification. Interestingly, instruments that were confused with each other led to non- overlapping regions and were confused when they were filtered in the most salient region of the other. These results suggest that musical instrument timbres are characterized by specific spectrotemporal modulations.

Email: etienne.thoret@mcgill.ca; philippe.depalle@mcgill.ca; stephen.mcadams@mcgill.ca

23

## Error correction neural correlates of auditory and visual sensorimotor synchronization

**Daniel Comstock*** *University of California, Merced*
**Ramesh Balasubramaniam** *University of California, Merced*

Inserting temporal perturbations into a metronome during a sensorimotor synchronization tapping task produces synchronization errors that are quickly corrected, even when the perturbation is too small to be perceived consciously. These error correction mechanisms were used in this study to investigate whether the timing mechanisms involved in sensorimotor synchronization tasks are universal or modality specific. EEG was measured during synchronization tapping tasks with either an auditory, or flashing visual metronome. Occasional temporal perturbations were inserted of +/- 60 *ms* and +/- 15 *ms* to produce period correction and phase correction responses, respectively. We found that while the auditory condition produced period corrections for the +/- 60 perturbations and phase corrections for the +/-15 *ms* perturbations, the visual condition produced only phase corrections for both the larger and smaller perturbations. The EEG data showed a clear Error Related Negativity (ERN) in the +60 auditory condition as expected, but no ERN was found for any of the visual conditions. However, the visual evoked EEG data showed a P1 component that was modulated by the direction of the +/- 60 conditions, and an increase in a P2 component modulated by the size of the perturbations, with the +/- 60 perturbations producing the largest increase in the P2 component. These findings suggest that the auditory system has privileged access to timing mechanisms in the brain, as even though the visual system was able to detect the perturbations, as evinced by the visual P1 and P2 ERPs, those detections did not translate into any kind of period correction response or an ERN as was shown in the auditory perturbations.

Email: dcomstock@ucmerced.edu; ramesh@ucmerced.edu

24

## Tonality without structure: Effects of tonal uncertainty and drone tones on emotional response to melodies

**Jeff Bostwick*** *University at Albany, SUNY*
**W. Neill** *University at Albany, SUNY*

Musical tonality (i.e., major vs. minor modes) has long been known to provoke emotional responses in listeners, but the mechanisms that underlie this effect are not well understood. One proposed explanation for the connection between minor modes and negative affect is that minor modes inherently cause more uncertainty in the listener, and this uncertainty is the source of the listener's negative affect. We tested this explanation by presenting subjects with several melodies played either unaccompanied or with one of seven droning bass notes to provide tonal context (i.e., certainty) and to encourage participants to interpret the melody as being presented in one of seven modes. Participants rated these melodies on affect (happy/sad) and musical tension. The melodies were intentionally created to be tonally ambiguous, with only the drone bass notes to provide context. Despite this ambiguity, we noted significant differences in affective response between melodies in the three major mode contexts as compared to those in the four minor mode contexts. These differences conformed to music theoretic expectations, such that major modes were rated as happier than minor modes. Contrary to the predictions of uncertainty theories, participants rated unaccompanied (i.e., less certain) melodies as significantly more positive and less tense than accompanied melodies. The results indicate that drone bass lines can establish a musical context, and that the absence of such a context does not lead to an increased perception of negative affect.

Email: jmbostwick@gmail.com; wneill@albany.edu

25

## The effects of tonal-harmonic structure on visual narrative processing

**Morgan Patrick*** *Brown University*
**Sheila Blumstein** *Brown University*
**Neil Cohn** *Tilburg University*
**John Mertus** *Brown University*

Schematic musical expectations facilitate cognitive processing when visual material is presented simultaneously with highly predictable musical chords. Previous research has debated whether this priming is specific to language or acts on domain-general processes. We suggest that adopting a narrative theory of tonal music can provide the key link between these claims via shared syntactic structure. Both Western tonal harmony and narrative syntax are hierarchical, recursive systems that unfold tension and release in time, suggesting a common processing mechanism during audiovisual integration. In this study, we tested for effects of expectancy cues across music and visual narrative by measuring online processing and recognition when chord progressions were synchronized to picture stories. Subjects were presented with 6-panel Peanuts comic strips whose constituents were either logically coherent or structurally and semantically unrelated (scrambled). Accompanying musical chord progressions were either tonal or atonal. Individual chords were aligned with picture onsets, and subjects monitored for picture targets in sequences presented panel by panel. Subjects were significantly faster to identify picture targets of comic strips when the musical accompaniment was tonal and the picture sequence was coherent, as compared to when music was atonal or pictures were scrambled. Furthermore, the magnitude of tonal facilitation was significantly greater for coherent picture stories than for scrambled, indicating an interaction between expectancy cues across the two modalities. Finally, highly salient musical phrase endings – cadences – facilitated online visual processing speed regardless of picture narrative coherence, providing empirical support for the attentional magnetism of musical phrase endings in multimedia contexts, even when visual scenes were semantically incoherent. No effect was found for musical expectancy on post-trial recognition. Implications for modality-general narrative processing mechanisms are discussed, with an eye toward framing tension-release trajectories in music as syntactic in nature.

Email: morgan_patrick@brown.edu; sheila_blumstein@brown.edu; neilcohn@visuallanguagelab.com

26

## Preference for colored audio noise

**Andrew Slifkin*** *Cleveland State University*
**Katrina Wolfgang** *Cleveland State University*

When the output of any system is measured over a long time period or event sequence, the structure of that output can be identified as a type of colored "noise." A wide variety of physical and biological systems have shown pink noise in their output, where pink noise is characterized by an inversely proportional relation between spectral power and frequency (1/f1). Voss and Clarke (1978) demonstrated a pink-noise structure for popular genres of Western music too. In addition, those authors assessed participants' preferences for computer-synthesized music with a white-noise structure—where spectral power remained constant over the range of frequencies (1/f0)—a pink-noise structure, and a brown-noise structure—where the fall off of power as a function of frequency was double (1/f2) that of pink noise. They reported that participants had the greatest preference for their synthesized pink-noise music. While Voss and Clarke provided the details of how they created their colored-noise music, they provided limited description of their psychophysical methodology and no numerical psychophysical results were reported. The current study was an attempt at a more formal, controlled assessment of preference for sound structure. Here, participants provided preference ratings for pure audio-noise samples. According to the results, as audio-noise structure changed from white to pink to brown noise, the group-mean preference ratings increased monotonically, i.e., brown noise was most preferred. Such results are in accord with more recent research showing that popular music may deviate from a pink-noise structure with a structure closer to brown noise.

Email: a.slifkin@csuohio.edu; katrina.wolfgang@gmail.com

27

## Comparing linguistic prediction in musicians and non-musicians

**Allison Fogel*** *Tufts University*
**Eddie Wlotko** *Tufts University*
**Gina Kuperberg** *Tufts University*
**Aniruddh Patel** *Tufts University*

Prediction or expectancy is increasingly thought to play an important role in the processing of both music and language, leading to recent interest in the possible relationship between cognitive mechanisms of prediction in the two domains. While there is evidence that individuals with musical training show enhancements in some aspects of language processing (such as prosody and affect), the impact of musical training on predictive tendencies in language has not been well explored. Because prediction is vital for successful music processing, it is possible that musical training may be associated with a greater tendency to predict upcoming information in general, or musical training may be associated with changes in other aspects of cognition (such as working memory) that in turn impact predictive tendencies in language. Wlotko, Federmeier, & Kutas (2012) have previously examined individual differences in lexical prediction strength in language as indexed by the amplitude of the ERP component known as the frontal positivity, finding considerable individual variation. Here, we use a similar sentence comprehension paradigm to explore differences in language prediction strength between individuals with and without musical training. Contrary to our predictions, no relationship was observed between frontal positivity amplitude (or any other ERP component) and musical training, suggesting that musical training does not impact the strength of an individual's lexical predictions in language. We will examine the results of cognitive testing of these participants, including measures of working memory and executive function, to explore if there is a relationship between any of these cognitive measures and an individual's prediction strength or degree of musical training.

Email: allison.fogel@tufts.edu; ewlotko@nmr.mgh.harvard.edu; gina.kuperberg@tufts.edu; a.patel@tufts.edu

28

## Systematic variation in rhythm production as tempo changes

**Scott Barton** *Worcester Polytechnic Institute*
**Laura Getz** *Villanova University*
**Michael Kubovy** *University of Virginia*

We investigated the effect of tempo on the production of the syncopated 3–2 son clave rhythm. We recorded eleven experienced percussionists performing the clave pattern at tempi ranging from 70 bpm to 210 bpm. As tempo increased, percussionists shortened the longest intervals and lengthened the shortest interval towards an intermediate interval that is located in the first and second positions in the pattern. This intermediate interval was stable across tempi. Contrary to prior studies, we found that the complexity of interval ratios had little effect on production accuracy or stability and the "short" interval in the pattern was not particularly stable. These results suggest that as tempo is varied, (1) experienced musicians systematically distort rhythmic intervals, (2) rhythmic configuration, and not just the complexity of interval ratios, affects the production of rhythmic intervals, and (3) the distinction between long and short intervals is context–dependent.

Email: sdbarton@wpi.edu; laura.getz@villanova.edu; kubovy@virginia.edu

29

## How do we decide when an auditory sequence ends and another begins?

**Lauren Amick*** *Michigan State University*
**Audrey Drotos** *Michigan State University*
**J. Devin McAuley** *Michigan State University*

Auditory rhythms yield the perception that elements of a sequence are connected in time. One perspective is that if the time interval between two auditory events is too long, then the second event is perceived to be isolated in time from the first. Here, we investigated the perception of 'connectedness' in time by having individuals make decisions about when one auditory sequence ends and another begins. Participants listened to a sequence of three tones and judged whether they heard a single group of three tones, or whether they heard a group of two tones following by an isolated tone. In all trials, the inter-onset-interval (IOI) between the first two tones defined a constant referent interval, T, while the IOI between the second and final tones varied across trials (T + $\Delta$T); $\Delta$T varied as a percentage of T between 0% and 500%. Sequences were presented using the method of constant stimuli. Results revealed that as the IOI between the second tone and the third tone was lengthened, participants, as expected, were increasingly likely to hear the final tone as starting a new sequence. However, when the interval marking the onset of the final tone occurred at multiples of the referent interval, participants tended to perceive the final tone as part of the group (i.e., they heard a group of three tones). This tendency to connect events in time at multiples of the referent interval resulted in an oscillatory pattern of responding that was a function of the referent interval. Results favor entrainment- over interval-based models of timing.

Email: laurenjamick@gmail.com; drotosau@msu.edu; dmcauley.msu@gmail.com