APCAM 2012

11th Annual Auditory Perception, Cognition, and Action Meeting

Thursday, November 15th

Minneapolis Hilton Hotel

Minneapolis, MN, USA

Program sponsored by

WASHBURN UNIVERSITY

Welcome to APCAM 2012

APCAM was founded on the notion " to bring together researchers from various theoretical perspectives to present focused research on auditory cognition, perception, and aurally guided action". Today, more than a decade after it began, APCAM remains one of the very few outlets where mixed within the same session you can will find a mixture of auditory researchers doing basic and applied work, exploring constructive and direct perception models, testing low-level and high-level accounts, as well as evaluating the processing of speech, music, and environmental noises.

An event like this cannot happen without needed support. Thanks are due to several groups who have made this year's APCAM possible, including the Psychonomic Society for covering room and equipment fees. A special thanks also is in order for Washburn University, which was responsible for the creation and printing of this conference program. These contributions have enabled this year's meeting to be offered free of charge to attendees. This is very unusual for conferences these days, and only seems fitting given the motivation to share scientific findings from different perspectives that really sets APCAM apart from other events.

Ultimately, perhaps the biggest debt is owed to the many regular APCAM attendees over the years who have helped to make this a great scientific forum. Additionally, thanks to the many newcomers in 2012 for helping to further enrich the meeting's content. We hope that, regardless of whether you are a regular or a newcomer, you will continue to include APCAM in your conference circuit in the future, and that you will feel free to invite colleagues to join us. Your contributions have helped deliver a great meeting for 2012. Have an enjoyable and productive day.

Sincerely,

The APCAM 2012 Organizing Committee
Michael D. Hall (Chair)
Mike Russell
Peter Q. Pfordresher
Devin McAuley
Kristopher Patten
Jeremy Gaston
John Neuhoff

APCAM 2012 Schedule

8:00	**Registration – Conrad D room at the Hilton Hotel**	
8:30	**Opening Remarks**	
	General (abstracts pages 7 – 8)	
8:40	Exploring the Role of Amplitude Envelope in Duration Estimation: Evidence for Two Strategies	**Michael Schutz* Guillaume Vallet David Shore**
9:00	Perceptual Discrimination of Listener-Target Position in Simulated Gunshot Sounds	**Michael Hall* Christopher Becker Jeremy Gaston Paul Fedele**
9:20	Auditory-Motor Priming Using Sounds Produced by Simple Actions	**Laurie Heller* Nicolas Zuniga-Penaranda Guillaume Lemaitre Nicole Navolio**
9:40	**Break (20 mins)**	
	Speech (abstracts pages 9 – 11)	
10:00	Sounding Like A Criminal: The Effect of Similarity and Voice Frequency on Earwitness Confidence Ratings	**Doug Alards-Tomalin* Jason Leboe-McGowan Launa Leboe-McGowan**
10:20	The Influence of Talker Variability on the Interpretation of Pragmatic Intent	**Jennifer Roche* Rick Dale**
10:40	An Approach for Manipulating the Perceived Gender of a Talker's Voice for Studies of Indexical Effects in Speech Perception	**Joe Toscano* Bob McMurray**
11:00	Modeling Speech Production with Bayesian Inference: Competition, Latency, and Articulation	**Christo Kirov* Colin Wilson**
11:20	Channeling the Cocktail Party Effect	**Kristopher Patten* Steven Holloway Michael McBeath**

Poster Session (11:50 AM – 1:20 PM) Minneapolis Convention Center – Ballroom A (Level 1) Abstracts located on pages 16 – 26		
Lunch (1:20 – 2:00 PM)		
Music (abstracts pages 12 – 13)		
2:00	**Keynote Speaker: Pitch Perception in Speech and Music**	**Andrew Oxenham***
2:30	Specificity of the Effects of Tone Language Experience on Melody Perception	**Evan Bradley***
2:50	Music Training and Cognitive Abilities: What Is the Direction of Causation?	**E. Glenn Schellenberg***
3:10	Musical Experience Effects on Perceptual Grouping	**Keturah Bixby* Elizabeth Marvin Joyce McDonough**
3:30	**Break (20 min)**	
Multi-modal (abstracts pages 14 – 15)		
3:50	Two Mechanisms of Sequence Perception When Performing Auditory Spectral Temporal Order Judgment	**Leah Fostick Dovrat Miron* Liron Tuval* Harvey Babkoff**
4:10	"Deafness" Effects in Detecting Alterations to Auditory Feedback During Sequence Production	**Peter Pfordresher***
4:30	"Clave Rap": The Segmentation of Linguistic Auditory Necklaces	**Laura Getz* Minhong Yu Priyanka Salona Michael Kubovy**
4:50	Effects of Velocity and Visual Stimuli on the Auditory Fröhlich Effect for Pitch	**Timothy Hubbard* Susan Ruppel**
5:10	**Closing Remarks**	

Posters (abstracts located on pages 16 – 26)		
1	Masked- and Unmasked Speech Effects Recollection of Semantically Categorized Words.	Anders Hurtig*
2	Tone Duration and Inter-Stimulus Interval Have Equal Effect on Dichotic Temporal Order Judgment	Leah Fostick* Harvey Babkoff
3	What is the Shape of the Sound Space Within Which We Perceive and Act? An Investigation Based on Euclidean and Action-Based Metrics.	Michael Russell*
4	Sound-Localization Performance and Sensitivity to Spectral Shape	Guillaume Andéol* Lionel Pellieux Ewan Macpherson Andrew Sabin
5	Grouping by Proximity and Grouping by Similarity in Auditory Necklaces	Minhong Yu* Michael Kubovy
6	The Scale Illusion Revisited: Hidden Melody Recognition and Musical Expertise	Nicholaus Brosowsky* Todd Mondor
7	The Effects of Language of Origin on Rhythm in Improvisational Jazz Solos	Harrison Lane Goldberg Claudia R. Thompson John G. Neuhoff*
8	Neighborhood Density and Grammatical Class: Within-Class Competitors Benefit Spoken Word Recognition	Julia Strand*
9	Delayed Auditory Feedback and Movement, Revisited: Bimanual Versus Unimanual Coordination	Peter Q. Pfordresher* Caroline Palmer Simone Dalla Bella Brian T. Kraus
10	Perceptual Evaluation of a Binaural Beamforming Algorithm	Jumana Harianawala* Susie Valentine Ivo Merks
11	Integration of Vestibular and Auditory Input in the Interpretation of Dynamic Sound Localization Cues	Janet Kim* Ewan Macpherson
12	Auditory Learning-Related Shifts in Generalization: A Case for Distortions of Representation	Matthew G. Wisniewski* Barbara A. Church Eduardo Mercado III

13	**Automatic Classification of Multi-Class Acoustic Environments**	**Martin F. McKinney* Anil Shankar Tao Zhang**
14	**Auditory Stimulus Generation Tools in MaxforLive**	**Michael D. Hall* Christopher Becker Thomas Redpath Ashley Assgari**
15	**Melodic Continuation In Three Dimensions: Comparing Expectation in Sequences of Pitch, Brightness, and Loudness**	**Jackson E. Graves* Christophe Micheyl Andrew J. Oxenham**
16	**Speaker Recognition with Pitch-Shifted Voices**	**Elan Barenholtz* Brittany Guidone**
17	**Interactions of Pitch and Timbre: How Changes in One Dimension Affect Discrimination of the Other**	**Emily Allen* Andrew Oxenham**
18	**Temporal Preparation and the Perceived Duration of Auditory Oddballs**	**Elisa Kim* Katherine B. Jones J. Devin McAuley**
19	**Neural Activity Relates to Melodic Interval Probability.**	**Roger E. Dumas* Arthur C. Leuthold Joshua K. Lynch Apostolos P. Georgopoulos**

8:40

Exploring the Role of Amplitude Envelope in Duration Estimation: Evidence for Two Strategies

Michael Schutz* *McMaster University*
Guillaume Vallet *Laval University/Lyon University*
David Shore *McMaster University*

Although many of the sounds we hear everyday are "percussive" in nature (i.e., produced by two objects impacting), perceptual experiments frequently use tones synthesized with "flat" or abruptly ending amplitude envelopes. These artificial sounds afford a specialized estimation strategy involving calculating the elapsed time between tone onset and offset—a strategy that would be problematic for the sounds common to our everyday environment that decay as they lose energy. In two experiments, we compared duration judgments for sounds with percussive (i.e. gradually decaying) and flat (i.e. abruptly ending) envelopes. Pairs of the same or mixed kinds of tones were presented in different blocks of trials in Experiment 1. More uncertainty was added in Experiment 2 by comparing a block of uniform kinds (flat-flat and percussive-percussive) with a block where all combinations were possible. The results showed similar sensitivity to judgments of the durations of flat and percussive sounds when participants could predict the envelope structure of the tones prior to their onset (allowing the opportunity to choose an optimal judgment strategy). However, participants demonstrated poorer performance for percussive sounds relative to flat when the trials were intermixed (forcing use of a single, general purpose strategy). The results suggest different strategies may be used when judging the durations of abruptly ending flat tones vs. gradually decaying percussive tones. We will discuss implications of these results in light of the pervasive nature of flat tones in auditory research, tones that may afford duration judgment strategies unavailable in everyday listening situations involving natural sounds.

Email: schutz@mcmaster.ca; gtvallet@gmail.com; dshore@mcmaster.ca

9:00

Perceptual Discrimination of Listener-Target Position in Simulated Gunshot Sounds

Michael Hall* *James Madison University*
Christopher Becker *James Madison University*
Jeremy Gaston *Army Research Laboratory*
Paul Fedele *Army Research Laboratory*

Gunshots provide potentially useful acoustic cues to the listener's position/angle relative to the target line, including differences in the relative intensity and timing of ballistic crack and muzzle blast components. Two experiments were conducted with 17 untrained listeners to determine if such listeners are sensitive to these cues. Both experiments involved simulated single-shot tokens that were based upon spectral and amplitude analyses from AK-47 samples. Sounds were generated using a MIDI-controllable synthesizer made in MaxforLive. Noise first was filtered to approximate spectral slope and limit maximum frequency, and then to generate spectral envelopes for ballistic crack and muzzle blast. Amplitude was updated each ms. Ballistic crack intensities and onsets relative to muzzle blasts were predicted for each angle (0-45 degrees in 5-degree increments) using a physical model developed at the Army Research Laboratory. In the AX discrimination task of Experiment 1 blocks of trials were distinguished by the standard on each trial (0, 25, or 45 degrees). In Experiment 2 listeners instead were asked which of two tokens on a trial represented a shot that passed closer to the listener. Listeners easily detected differences in firing angle, with mean sensitivity (d') in Experiment 1 approaching ceiling. Sensitivity as a function of angle also differed across standards, consistent with the larger observed acoustic differences as angle approached 0 degrees. Experiment 2 revealed that listeners treated the 0-degree stimulus as a perceptual anchor, with accuracy remaining relatively high independent of angle only for that standard. In the remaining conditions accuracy reliably exceeded chance only for pairings with the 0-degree token. The reported findings indicate that it should be possible to train soldiers to accurately recognize their position relative to the target line. Implications for training procedures and suggested future research concerning these cues will be discussed. [supported by DCS Corp. subcontract APX03-S002]

Email: hallmd@jmu.edu; chris.becker4@gmail.com; jeremy.r.gaston.civ@mail.mil; paul.d.fedele.civ@mail.mil

9:20

Auditory-Motor Priming Using Sounds Produced by Simple Actions

Laurie Heller* *Carnegie Mellon University*
Nicolas Zuniga-Penaranda *Carnegie Mellon University*
Guillaume Lemaitre *Università Iuav di Venezia, Venice, Italy*
Nicole Navolio *Carnegie Mellon University*

We explored a new form of auditory-motor priming. Motor priming exists if an action is performed more rapidly after the presentation of facilitating cues than after the presentation of interfering cues. We hypothesized that environmental sounds could be used as cues to create motor priming. To create facilitation, we devised a congruent priming sound that was similar to the sound that would be made by the gesture that was about to be performed. To create interference, we devised an incongruent sound that would not normally be made by the gesture that was about to be performed. Using this paradigm we found evidence of auditory-motor priming between environmental sounds and simple gestures such as tapping and scraping. First, participants were initially familiarized with making two different gestures with an interface. The gestures were actions, such as tapping and scraping, that normally make a sound in the environment. Participants were taught to respond to a directional command (such as left/right or up/down) to make one of two possible motor responses (actions). In some conditions, while the subjects learned the actions they also heard the normally occurring ("natural") sounds as they were produced, but in other conditions the sound-action relationship was altered (either by silencing a sound or associating an arbitrary sound). Next, participants were tested on how quickly they could make the correct motor response to a directional command. Sometimes the directional command was preceded by a sound cue but sometimes the sound was absent. The sound cue was equally likely to be congruent or incongruent with the gesture that was about to be performed. We predicted that motor responses would be faster and more accurate when they were preceded by congruent sounds than when they were preceded by either incongruent sounds or no sound cue. We found evidence for auditory-motor priming over a range of conditions: priming occurred both when two hands were used and when only one hand was used; priming also occurred both when the actions involved manipulating wooden objects and when the interface was a computer keyboard; finally, priming occurred both when the sound cues were consistent with naturally occurring sounds and when the sound cues were arbitrarily associated to the actions. Thus we found evidence for relatively rapid learning of action-sound pairings even if they were arbitrary.

Email: hellerl@andrew.cmu.edu; nzunigap@andrew.cmu.edu; glemaitre@iuav.it; nvn@andrew.cmu.edu

10:00

Sounding Like A Criminal: The Effect of Similarity and Voice Frequency on Earwitness Confidence Ratings

Doug Alards-Tomalin* *University of Manitoba*
Jason Leboe-McGowan *University of Manitoba*
Launa Leboe-McGowan *University of Manitoba*

The current study examined the previously found deep voice heuristic (Alards-Tomalin et al., 2011), a tendency for participants to be more confident that a voice "line-up" suspect was a previously heard criminal in a simulated robbery, when the suspect's voice was modified to be lower in pitch versus higher. Additionally, the unmodified suspect voice received the highest overall confidence ratings over when it was modified to be high or low in pitch. In the current experiment we were interested in determining if confidence ratings are based on matching the similarity of the unmodified suspect voice to the original criminal, or were being driven by matching a perceptual quality of the unmodified sound files to the original unmodified criminal recording. To test this, the lineup contained both suspects and the criminal voice. If ratings are based on the quality of the recording, than there should be no significant difference in the confidence ratings ascribed to the unmodified suspect and criminal voices in the lineup. If confidence ratings are being driven by a memory-based assessment regarding the similarity of the lineup voices, the criminal voice should receive the highest confidence ratings followed by unmodified suspects. Participants exhibited the highest confidence in the criminal voices, followed by unmodified suspects, followed by the low pitch suspects. Lastly, high pitched suspects were rated with the lowest confidence. These results indicate that confidence is driven by the degree of similarity of the suspect's vocal attributes to the original criminal voice. Additionally, the results also replicate the earlier finding that low-frequency modified voices are perceived to be more criminal in nature – receiving significantly higher confidence ratings than the same suspect voice manipulated to be higher in frequency.

Email: umalards@cc.umanitoba.ca; leboej@cc.umanitoba.ca; leboe@cc.umanitoba.ca

10:20

The Influence of Talker Variability on the Interpretation of Pragmatic Intent

Jennifer Roche* *University of Rochester*
Rick Dale *University of California, Merced*

Language and prosodic cues to affective speech are essential for successful interpretation of intent. Two experiments (production and perception) evaluate the impact of affective prosody on the processing and interpretation of affectively spoken language. We investigate the role of affective prosody in the production and perception of four affective categories (neutral, compassion, irritation and sexual innuendo). We specifically examined the acoustics of these categories between and within four native speakers of Southern-American English (2 males, 2 females). Results suggest that while there was evidence of systematic acoustic correlates between and within talkers, affect production involves a substantial amount of talker variability. For example, Compassion and Innuendo ranged in F0 between talkers (at least $p < .05$), but Compassion and Innuendo was marked by higher F0 relative to Irritation and Neutral expressions for all talkers (at least $p < .05$). In a second experiment, we were able to preview real-time processing of intent by utilizing a new method of tracking responses semi-continuously during affective speech perception via computer-mouse movements (x & y mouse cursor coordinates). Results from the perception study suggest that despite talker variability, participants (N = 24) seemed to accurately categorize intent (accuracy ~70%). However, confidence measures revealed that affective category and talker influenced responding, despite high accuracy in affective intent. For example, listeners had an increasingly difficult time responding accurately to Compassion expressions overall, especially for one specific talker. When Compassion was accurately categorized, arm trajectory measures showed that listeners were more hesitant, showing less confidence in their response (e.g., increased x-flips and longer reaction time; at least $p < .05$). Thus, it would seem that talker variability in the form of prosodic variation influences a listener's ability to quickly integrate the affective information, as well as the confidence in their interpretation during the online processing of intent.

Email: jroche@bcs.rochester.edu; rdale@ucmerced.edu

10:40

An Approach for Manipulating the Perceived Gender of a Talker's Voice for Studies of Indexical Effects in Speech Perception

Joe Toscano* *University of Illinois*
Bob McMurray *University of Iowa*

A number of researchers have become interested in the effects of indexical information on speech perception. A commonly studied indexical cue is talker gender, which produces systematic changes in acoustic cues that are relevant for speech. However, it has been difficult to manipulate listeners' perception of gender, making controlled experiments examining its effects difficult to run. Here, we present a novel approach for manipulating the perceived gender of a talker's voice that involves changing fundamental frequency, formant frequencies, and formant bandwidths. We generated stimuli using this approach and demonstrate that listeners correctly categorize the sounds at both endpoints of a male/female continuum, approximating the likelihood of male and female responses seen with natural stimuli. In contrast, a simpler method that involves only manipulating fundamental frequency does not produce correct responses at both ends of the continuum. We also used this approach to create stimuli that had an ambiguous gender and found that preceding talker context (i.e., whether the preceding sentence was spoken by a man or woman) had an effect on listeners' categorization of the stimuli. Thus, the approach presented here provides a way for researchers to manipulate perceived talker gender in order to examine the effects of contextual variation on speech perception.

Email: jtoscano@illinois.edu; bob-mcmurray@uiowa.edu

11:00

Modeling Speech Production with Bayesian Inference: Competition, Latency, and Articulation

Christo Kirov* *Johns Hopkins University*
Colin Wilson *Johns Hopkins University*

When words are elicited in the context of minimally different neighbors, speakers produce them with phonetic hyperarticulation and longer onset latencies. These effects both attenuate rapidly as the feature distance between the word and its neighbor increases (e.g., cap shows hyperarticulation/longer latency in the context of gap and tap, with which it differs only in voicing or place of articulation, but not nap, with which it differs in at least voicing, place of articulation, and nasality). We show that the link between hyperarticulation and latency follows from a model in which speech production involves incremental Bayesian inference at multiple levels of processing. Lexical selection processes send messages to phonology-level modules indicating which phonological forms should be produced. Inference is performed each time a new message is received, altering the distribution of proposed forms. Due to noise in the production system messages are uncertain, and multiple similar forms may be partially activated by a message. This results in implicit competition and a longer time (i.e., more messages received) until one form reaches a high enough probability to be selected for production. While a phonological form is being selected, the phonological selection process also sends messages to phonetic modules indicating how phonetic parameters should be set. Again, distributions over phonetic parameters are altered by Bayesian inference. When phonetic parameters are initialized with a strong prior favoring less effortful articulations, the time it takes to select a phonological form for production is correlated with the total amount of hyperarticulation in the phonetic realization of the form. We apply the model to novel and previously reported experiments on VOT duration (Baese-Berke and Goldrick, 2009) and previously reported results on primed speech production (Meyer & Gordon, 1985).

Email: kirov@cogsci.jhu.edu; wilson@cogsci.jhu.edu

11:20

Channeling the Cocktail Party Effect

Kristopher Patten* *Arizona State University*
Steven Holloway *Arizona State University*
Michael McBeath *Arizona State University*

Numerous studies have investigated the phenomenon of a salient, yet unattended auditory stimulus (i.e., a participant's given name) cutting into concentration of an attended channel that is otherwise stable. Cherry (1953) called this the Cocktail Party Problem and in recent years this phenomenon has been referred to the Cocktail Party Effect. To date, studies have focused primarily on presenting the target and distractor channels binaurally. The current study uses both binaural and monaural presentation to investigate the role of mechanical (stapedius innervation) and neural inhibition. In addition, a psychophysical paradigm was assessed for its efficacy in replicating the results from traditional methods of cocktail party stimulus presentation and EEG signals were recorded to examine brain wave activity as a function of the task. Results show that when both target and distractor (with name) channels are presented to the same ear, participants detect their name with only 13% accuracy. This is significantly less than when the channels are presented to opposite ears (86% hit rate).

Email: kjp@asu.edu; srh@asu.edu; m.m@asu.edu

2:00

Keynote Speaker: Pitch Perception in Speech and Music

Andrew Oxenham* *University of Minnesota*

Pitch is a primary auditory percept, but many aspects of its coding remain unclear. Numerous theories and models for how pitch is extracted by the auditory system have been proposed over the past century or more, and yet the topic remains controversial. This talk will review some basic theories of pitch, and will highlight some of the findings from our lab, which shed light on how pitch is represented in the auditory system. In the first part, we will review the relationship between temporal fine structure and temporal envelope and their relative contributions to the perception of speech and music. Temporal fine structure is currently thought to play a major role, not just in music, but also in understanding speech in noisy or complex acoustic environments. In particular, specific deficits in the perception of temporal fine structure have been proposed to explain some of the communication challenges faced by people with hearing loss and cochlear implants. Recent data from our lab question whether temporal fine structure is coded temporally, or whether such deficits underlie speech communication challenges. In the second part, we will outline some recent evidence suggesting that the relationship between musical pitch perception and auditory temporal coding may not be as clear as previously thought. Finally, we will present new data on the perception of whispered speech in various acoustic backgrounds, to argue that deficits found in hearing-impaired listeners and cochlear-implant users, which had previously been ascribed to a loss of pitch information, may in fact be more readily explained in terms of changes in spectral resolution. [Work supported by NIH grant R01 DC 05216.]

Email: oxenham@umn.edu

2:30

Specificity of the Effects of Tone Language Experience on Melody Perception

Evan Bradley* *Lawrence University*

This study investigated whether the advantages of tone language speakers in melody perception versus nontone language speakers can be linked to specific properties of their tone systems. If so, this suggests that melody and lexical tone perception rely on shared general auditory resources, consistent with perceptual learning models (Ahissar, 2009; Patel, 2009; Patel, 2011). Important acoustic cues to lexical tone (pitch height, direction, slope) are argued to correspond to properties of musical melody (key, contour, interval). Languages rely on these to varying degrees, so speakers of different languages were hypothesized to vary in sensitivity to melody as a function of tonal properties of their language. English, Yoruba, and Mandarin-speaking nonmusicians took a melody discrimination test, detecting changes to the key, contour, or intervals of melodies. Direction and slope are important cues to Mandarin tone perception, so Mandarin speakers were expected to be more sensitive than English speakers to contour and interval. Direction is important to Yoruba tone perception, while slope is of moderate importance, so Yoruba speakers were expected to be more sensitive than English speakers to contour and interval, though less so on interval than Mandarin speakers. Mandarin and Yoruba listeners outperformed English speakers in perception of interval, but did not differ from one another. Only Mandarin speakers outperformed English speakers on contour. The three groups did not differ in discrimination of key. Findings were consistent with the general hypothesis that melody perception associated with tone language is specific to certain properties of melody. Yoruba listeners' performance only partially matched expectations, indicating that assumptions about Yoruba tones should be reexamined, and/or the mapping between tonal and melodic pitch properties revised. However, Yoruba and Mandarin speakers did not perform identically, supporting the hypothesis that each tone language influences melody perception differently.

Email: evan.d.bradley@lawrence.edu

2:50

Music Training and Cognitive Abilities: What is the Direction of Causation?

E. Glenn Schellenberg* *University of Toronto*

Over the past 20 years, much research has examined the possibility of associations between music lessons and cognitive abilities. Music training is now known to be associated with enhanced performance on tests of hearing, speech perception, attention, phonological awareness, vocabulary, reading, verbal memory, nonverbal memory, working memory, visuospatial abilities, mathematical reasoning, academic achievement, and IQ. Naturally, these behavioral differences are accompanied by differences in brain structure and function. Although the vast bulk of the relevant research has used quasi-experimental or correlational methods--which preclude inferences of causation--the general belief is that music training causes the observed benefits. Some scholars even claim that music training is the ideal model for the study of plasticity. The goal of the present paper is to challenge this prevailing view, suggesting instead that pre-existing individual differences influence (1) who takes music lessons, and (2) performance on most tests of cognitive abilities whether they measure low-level listening skills or IQ.

Email: g.schellenberg@utoronto.ca

3:10

Musical Experience Effects on Perceptual Grouping

Keturah Bixby* *University of Rochester*
Elizabeth Marvin *University of Rochester, Eastman School of Music*
Joyce McDonough *University of Rochester*

Auditory perceptual grouping occurs when the brain imposes a grouping organization on auditory sequences. It has been observed that perceptual grouping follows the iamb-trochee law, meaning that sounds alternating in duration group as iambs (short-long), while those alternating in amplitude group as trochees (loud-soft), and this serves as a basis for linguistic metrical structure (Hayes, 1996). However, Iversen, Patel and Ohgushi (2008) found duration sequences are not consistently grouped as iambs by native Japanese speakers, whose language allows an unusual metrical structure. We asked whether musical experience can similarly affect perceptual grouping. Following the method of Iversen et al. (2008), participants listened to tone sequences alternating in duration or amplitude, and reported their perceived grouping. The comparison groups were people with equivalent years of musical experience enrolled in a university (n=43) or in a conservatory (n=24), and people in the university who had little to no musical experience (n=32). For the analysis, we performed a regression using an empirical logit transform, with participant group as the key predictor of interest. All musicians grouped alternating amplitudes as loud-soft significantly more than nonmusicians, which was likely due to the increased salience of subtle amplitude cues given musical experience. Conservatory musicians grouped alternating durations as short-long significantly less frequently than all other groups, including the university musicians. We hypothesized that conservatory students' formal musical instruction through theory courses could account for their difference from other musicians with similar years of experience. We tested a fourth group of university theory students (n=18) and found that they grouped duration sequences no differently than the conservatory musicians. This hints that differences in duration grouping between musician groups are due to formal training, not just musical experience. Our results suggest auditory perceptual grouping is influenced by rhythmic experience, especially from music practice and study.

Email: kbixby@bcs.rochester.edu; bmarvin@esm.rochester.edu; joyce.mcdonough@rochester.edu

3:50

Two Mechanisms of Sequence Perception When Performing Auditory Spectral Temporal Order Judgment

Leah Fostick* *Ariel University Center*
Dovrat Miron *Hadassah Academic College*
Liron Tuval *Hadassah Academic College*
Harvey Babkoff *Ashkelon Academic Center*

Two types of sequence perception were previously suggested: (1) direct perception of the elements in the pair of tones and then judgment of the order of their occurrence; (2) holistic perception of the tonal patterns created by the order of the elements without separate perception of the elements. We suggested previously that the dichotic temporal order judgment (TOJ) paradigms (two brief tones that are presented to different ears) are successfully performed by the direct-temporal perception; while the spectrally-based TOJ paradigms (two brief tones with different frequencies), are successfully performed by the holistic perception. In the current study we show that spectral TOJ can be performed by both types of perception, depending on the difference in frequency between the tones to be judged. We posited that TOJ thresholds of 30-60 msec (the published range of TOJ threshold) would indicate direct perception, while a TOJ threshold of around 0 msec would reflect holistic perception. Sixty-three participants were divided randomly into three groups and performed spectral TOJ with one out of three frequency combination conditions: (1) 1,000-1,100Hz; (2) 1,000-2,000Hz; and (3) 1,000-3,500Hz. More participants in the 1,000-3,500Hz condition (14%) had TOJ thresholds within the expected range (30 – 60 msec) than in the 1,000-1,100Hz (5%) and the 1,000-2,000Hz (5%) conditions. On the other hand, more participants in the 1,000-1,100Hz (40%) and the 1,000-2,000Hz (48%) conditions had TOJ thresholds of 0 msec than in the 1,000-3,500Hz condition (32%). The results indicate that when the tone frequencies are closer, participants tend to judge their order by perceiving their global envelop, rather than each one of them separately as they tend to do when the frequencies are more dispersed. These results support the view of previous researchers of two types of sequence perception, and support the claim of two distinct mechanisms for performing TOJ.

Email: leah.fostick@ariel.ac.il; dovratm14@gmail.com; lirontuval@yahoo.com; babkofh@gmail.com

4:10

"Deafness" Effects in Detecting Alterations to Auditory Feedback During Sequence Production

Peter Pfordresher* *University at Buffalo, The State University of New York*

Past research has shown that when discrete responses are associated with a perceptual goal, performers may have difficulty detecting events that are commensurate with that goal. Three experiments are reported here that test whether such effects extend to sequence production. In Experiment 1, participants performed 8-note melodies repeatedly, and on each trial a single tone could be altered with respect to its pitch and/or synchrony with actions. Results suggested a selective deficit of detection when feedback pitch was unchanged, but the event was slightly delayed. Experiment 2 showed that this "deafness" to feedback is limited to motor tasks that require sequencing, in that similar effects did not emerge when participants produced pitch sequences by tapping a single key repeatedly. A third experiment demonstrated similar results to Experiment 1 when the mapping of keys to pitches on the keyboard was reversed. Taken together, results suggest a selective deafness to response-congruent delayed feedback, consistent with the idea that performers suppress previously planned events during production.

Email: pqp@buffalo.edu

4:30

"Clave Rap": The Segmentation of Linguistic Auditory Necklaces

Laura Getz* *University of Virginia*
Minhong Yu *University of Virginia*
Priyanka Salona *University of Virginia*
Michael Kubovy *University of Virginia*

We investigated the intersection of rhythmic and syntactic processing by studying the segmentation of linguistic auditory necklaces (cyclically repeating rhythmic patterns—ANs). In our previous work, we have found that two principles—the run and the gap principles—additively determine how listeners segment patterns. Here, we examined whether these grouping principles were stronger than syntactic segmentation. Our ANs varied with respect to rhythm and content: the rhythm was either a sixteen-beat five-note syncopated rhythm (…1001001000101000…) or a sixteen-beat five-note unsyncopated rhythm (…1010100000101000…), and the content was (a) a single percussive sound, (b) a single word, or (c) five different words that formed a sentence. On each trial, as soon as the AN began, a circular array of five icons appeared on the screen. Each icon was highlighted in clockwise order when a sound of the pattern played. Participants were instructed to determine which of the sounds seemed to be the starting point (the clasp) of the AN. To respond, they clicked on the icon corresponding to the clasp. In the nonlinguistic and the repeated-word ANs, the run and gap principles predicted the location of the clasp (as in previous studies). In the sentence ANs, the clasp most often was at the beginning of the sentence, although participants did occasionally follow the run and gap principles. This is Stroop-like effect: listeners found it hard to ignore the sentence and just respond to the rhythm. We will discuss these results in light of the progress that has been made in understanding the relations between music and language (see Patel, 2008).

Email: lmg5ep@virginia.edu

4:50

Effects of Velocity and Visual Stimuli on the Auditory Fröhlich Effect for Pitch

Timothy Hubbard* *Texas Christian University*
Susan Ruppel *University of South Carolina, Upstate*

Memory for initial pitch of an auditory target that increased or decreased in frequency (i.e., perceived as "ascending" or "descending", respectively, in pitch) was examined, and effects of pitch velocity and of whether onset of pitch motion was accompanied by a visual stimulus were considered. In Experiment 1, memory for initial pitch was displaced in the direction of pitch motion, and displacement was larger with faster pitch velocity. In Experiment 2, a visual stimulus briefly appeared before, at the moment of, or after pitch motion onset. Memory for initial pitch was displaced in the direction of pitch motion, but displacement was not influenced by whether the visual stimulus was presented before or after pitch motion onset. In Experiment 3, a visual stimulus appeared before, at the moment of, or after pitch motion onset, and the visual stimulus was visible throughout the duration of pitch motion. Memory for initial pitch was displaced in the direction of pitch motion, and there was a trend for larger displacement when the visual stimulus was presented before pitch motion onset. Forward displacement in remembered pitch in all experiments is consistent with a Fröhlich effect in memory for auditory pitch. Effects of the visual stimulus on the Fröhlich effect for pitch are generally consistent with effects of an auditory stimulus on visual representation momentum (i.e., effects of an auditory stimulus on memory for final location of a moving target) reported by Teramoto, Hidaka, Gyoba, and Suzuki (2010). Also, in all experiments, there was larger displacement when pitch motion descended than when pitch motion ascended, and this is consistent with representational gravity in memory for auditory pitch. Overall, memory for auditory targets in frequency space exhibited spatial biases similar to memory for visual targets in physical space.

Email: timothyleehubbard@gmail.com; sruppel@uscupstate.edu

Poster session (11:50 AM – 1:20 PM)
Minneapolis Convention Center – Ballroom A (Level 1)

1

Masked- and Unmasked Speech Effects Recollection of Semantically Categorized Words.

Anders Hurtig* *University of Gavle*

This research examined the effects of masked- and unmasked speech on the recollection of semantically categorized words. The words used in the experiment were sorted by semantic categories such as fruits, car brands and clothing. In reproducing the words an audio loop were played back through the headphones with words belonging to the same semantic category as the rehearsed words. The audio loop was presented eight times with the speech masked and with a broadband noise and eight times without a mask. The results showed that recall of masked speech increased the number of correct words in comparison to the number of correctly reproduced words in non-masked speech. It also showed that intrusions decreased in masked speech compared to intrusions in non-masked speech.

Email: anders.hurtig@hig.se

2

Tone Duration and Inter-Stimulus Interval Have Equal Effect On Dichotic Temporal Order Judgment

Leah Fostick* *Ariel University Center*
Harvey Babkoff *Ashkelon Academic Center*

The dichotic temporal order judgment (TOJ) paradigm has been used widely. However, the reported estimates of thresholds expressed as the inter-stimulus interval (ISI) for 75% accuracy are varied. In the current study, we examined the influence of tone duration, within the 10-40 msec range, on dichotic TOJ accuracy and threshold. Sixty-five participants were randomly divided into four groups, each performed dichotic TOJ with tone duration of 10, 20, 30, or 40 msec. When the proportion of left-leading responses (dichotic TOJ accuracy) is plotted as a linear function of ISI, PSE = 0 msec and R2=88%. When plotting the proportion of left-leading responses (dichotic TOJ accuracy) as a linear function of stimulus-onset asynchrony (SOA), PSE = 0 msec, but R2=94% (Fisher r-to-z transformation, z=-2.02, p=.02). ISI threshold decreases as a function of stimulus duration at a rate of approximately 8.6 msec per 10 msec increase in tone duration. When the group mean data were tested against a model that predicted a 10 msec reduction in threshold per decade increase in duration, the data did not deviate significantly (Probit analysis, Z=2.84; p=0.01). When plotting SOA threshold data as a function of tone duration, the four duration-group averages fall very close or on the zero slope line. The point at which the average dichotic TOJ thresholds (SOA) crosses the vertical axis is approximately 56.78 +/- 19.64 msec (Probit Analysis, Z=8.09, p=.001). The results suggest that ISI and tone duration affect dichotic TOJ thresholds equally, and when converted to stimulus onset asynchrony (SOA), are invariant to changes in tone duration within the range of 10-40 msec.

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

3

What is The Shape of the Sound Space Within Which We Perceive and Act? An Investigation Based on Euclidean and Action-Based Metrics.

Michael K. Russell* *Washburn University*

Based on the result of numerous investigations, it appears that auditory perception of space is nonlinear in nature. It has been known for quite some time that judgments of sound source distance tend to be overestimated for relatively nearby targets and underestimated for more distant targets. With respect to the perceived location of an unseen sound source, perceptual precision and stability changes significantly with the position of the auditory target. More specifically, perceptual accuracy decreases and perceptual variability increases as the target moves from the observer's midline to the periphery. Such findings suggest that the space within which organisms perceive and act is a nonlinear space. However, that conclusion may be erroneous and may simply be a reflection of the methods used in those investigations. For example, distance studies tend to involve targets at great distances, and distance and localization studies tend to require participants to make Euclidean judgments. The current study sought to determine the shape of sound space for nearby targets. The impact of Euclidean versus action-based metrics was also determined. The findings will be discussed in relation to the manner with which investigations into auditory spatial perception are conducted. The findings will also reveal the degree of similarity between auditory and visual perceptions of space.

Email: mike.russell@washburn.edu

4

Sound-Localization Performance and Sensitivity to Spectral Shape

Guillaume Andeol* *IRBA*
Lionel Pellieux *IRBA*
Ewan Macpherson *University of Western Ontario*
Andrew Sabin *Northwestern University*

Individual differences exist in sound-localization performance even for normal-hearing listeners. Some of these differences might be related to acoustical differences in localization cues carried by the HRTFs (Head Related Transfer Functions). For instance, among a group of listeners, HRTFs differ in their "spectral richness", the strength and diversity of spectral cues: a low spectral richness might limit sound-localization performance. Recent data suggest that individual differences in sound-localization performance could also have perceptual origins. The localization of an auditory target in the up/down and front/back dimensions requires the analysis of the spectral shape of the stimulus. In the present study, we investigated the hypothesis that individual sensitivity to spectral shape can account for some of individual differences in sound-localization performance. The sensitivity to spectral shape was evaluated by the spectral modulation detection thresholds assessed for different carrier bandwidths and spectral modulation frequencies. The spectral richness and the spectral modulation detection thresholds were measured in 2 similar groups of young naïve normal-hearing listeners. One group performed an absolute sound-localization task in silence with virtual auditory targets synthesized using individual HRTFs. The other group performed a forced-choice (8 alternatives) sound-localization task in noise with real auditory targets. No correlation between spectral modulation detection thresholds and sound-localization performance was found in the first group whereas significant correlations were found in the second group: the lower the thresholds for low spectral modulation frequencies, the better the sound-localization performance was. No correlation between sound-localization performance and spectral richness was found in any case. Those results suggest that perceptual-ability differences rather than acoustical differences contribute to individual differences in sound localization performance.

Email: guillaume.andeol@gmail.com; lionel.pellieux@irba.fr; ewan.macpherson@nca.uwo.ca; a-sabin@northwestern.edu

5

Grouping by Proximity and Grouping by Similarity in Auditory Necklaces

Minhong Yu* *University of Virginia*
Michael Kubovy *University of Virginia*

The two classic Gestalt grouping principles — grouping by proximity and grouping by similarity — have been studied intensively in visual modality. Here we used non-metric auditory necklaces (which are cyclic auditory patterns, AN in short) to investigate auditory grouping by temporal proximity and grouping by similarity of pitch. To quantify grouping by proximity, we manipulated the SOAs between two sets of tones (e.g. ... di - di - - di - di - - ...). To quantify similarity, we manipulated the frequency differences between two sets of tones (e.g. ... di - di - da - da - ...). We appied pairs of grouping principles conjointly to the same stimuli to make it multistable (e.g. ... di - da - - da - di - - di - da - - da - di - - ...), and asked whether their conjoint effect was additive (as has been found in visual grouping). On each trial, as soon as an AN was played, a circular array of squares (whose number corresponded to the number of tones in the AN) appeared on the screen. While each tone was played, a corresponding square was highlighted. The squares were highlighted in clockwise order. The participants' task was to click on the square corresponding to the tone they perceived as the starting point of the AN. They were not required to click on a square when it was highlighted.

Email: minhongyu@virginia.edu; kubovy@virginia.edu

6

The Scale Illusion Revisited: Hidden Melody Recognition and Musical Expertise

Nicholaus Brosowsky* *University of Manitoba*
Todd Mondor *University of Manitoba*

The current study was designed to investigate the magnitude of an auditory illusion known as the scale illusion (Deutsch, 1974, 1975) using a hidden melody recognition paradigm. The scale illusion occurs when listeners misperceive a complex melody consisting of two dichotic note patterns constructed from two interleaved diatonic major scales (one ascending and one descending). People typically perceive two wave-like melodies separated by frequency rather than the veridical perception based on ear of input. The magnitude of this misperception has been shown to vary considerably according to the data collection methods such as verbal reports and forced-choice questionnaires used in previous studies (e.g., Smith, Hausfeld, Power, & Gorta, 1982). The hidden melody recognition paradigm used in the present study was designed to eliminate the inconsistencies associated with these data collection methods. In each of three experiments, participants were presented with a hidden melody recognition test that assesses the listener's ability to identify simple tone patterns (upper and lower wave melodies, ascending and descending diatonic scales) within complex melodies. The experiment presents a replication of Deutsch's original scale illusion study measuring perceptions of note organizations based on frequency and good continuity under the scale illusion stimulus. Results of this study provide evidence for the utility of the hidden melody perception in measuring melody perceptions and note some fundamental differences between musician and non-musician perceptions of the scale illusion.

Email: umbroson@cc.umanitoba.ca; Todd.Mondor@ad.umanitoba.ca

7

The Effects of Language of Origin on Rhythm in Improvisational Jazz Solos

Harrison Lane Goldberg *The College of Wooster*
Claudia R. Thompson *The College of Wooster*
John G. Neuhoff* *The College of Wooster*

The rhythmic patterns of American and French improvisational jazz solos were compared to examine the potential influence of the language of origin, American English or French, on the performers' rhythmic improvisational styles. nPVI values for transcriptions of American and French jazz performances (50 each) were calculated. The mean nPVI for American transcriptions (M = 66.99) was significantly higher than the mean nPVI for French transcriptions (M = 49.27). These results parallel rhythmic differences between spoken English and French (Ramus, 2002) as well as between English and French classical music (Patel & Daniele, 2003).

Email: hgoldberg12@wooster.edu; crt@wooster.edu; JNEUHOFF@WOOSTER.EDU

8

Neighborhood Density and Grammatical Class: Within-Class Competitors Benefit Spoken Word Recognition

Julia Strand* *Carleton College*

Activation-competition models of spoken word recognition propose that acoustic-phonetic input activates a set of candidate lexical representations in memory, and that these representations then compete for recognition. Despite evidence from neuroimaging and lesion studies demonstrating that a word's grammatical class (e.g., noun, verb) influences how it is processed (Damasio & Tranel, 1993; Perani et al., 1999) models of word recognition have not yet included grammatical class as a variable influencing lexical competition.

Lexical decision times and naming times for 745 words from an existing database were analyzed (Luce & Pisoni, 1998). For each word, a measure of lexical competition, Frequency-Weighted Phi-Square Density (FWPSD), was calculated. FWPSD quantifies the perceptual similarity (and therefore, amount of competition) between a stimulus word and its competitors, weighted by their frequency of occurrence (Strand & Sommers, 2011). For each word, I also determined the proportion of FWPSD that comes from within-class competitors (e.g., for a noun stimulus word, the proportion of FWPSD that comes from noun competitors).

Regression analyses revealed that beyond the influence of word frequency (Brysbaert, New, & Keuleers, 2012), phonotactic probability (Vitevitch & Luce, 2004), and FWPSD, the proportion of FWPSD from within-class competitors explained a small but significant amount of variance in lexical decision times and naming times. Words that have a large proportion of their FWPSD from within-class competitors were recognized more quickly than those with a low proportion of competition from within-class competitors. These data could be explained with the existence of a grammatical class level of analysis in models of word recognition. Activation would spread bi-directionally from word-level representations to grammatical-class representations. Therefore, words that share grammatical class with many competitors would benefit from additional activation.

Email: jstrand@carleton.edu

9

Delayed Auditory Feedback and Movement, Revisited: Bimanual Versus Unimanual Coordination

Peter Q. Pfordresher* *University at Buffalo, The State University of New York*
Caroline Palmer *McGill University*
Simone Dalla Bella *WSFiZ in Warsaw*
Brian T. Kraus *University at Buffalo, The State University of New York*

Pfordresher and Dalla Bella (2011, JEP:HPP) reported two experiments concerning the effects of Delayed Auditory Feedback (DAF) on finger movements in isochronous, unimanual tapping. Analyses of motion data suggested that DAF affects the movement trajectory at the time of feedback onset. Whereas the finger's downward velocity predicted the effect of DAF on tapping rate, the finger's height at the time of the feedback onset predicted DAF's effect on timing variability. The current experiment adopted a similar approach in analyses of unimanual and bimanual tapping to address a new question: Is the effect of DAF best predicted by the movement state of finger used to generate the feedback tone (a feedback related effect) or the finger being prepared for the next tap (a movement planning effect)? Participants tapped isochronously either by alternating taps between the left and right index fingers (bimanual coordination), or by tapping repeatedly with the index finger of the same hand (unimanual coordination), while experiencing delays of different durations (in addition to a normal auditory feedback condition). Analyses of tapping data suggest that the coordination condition influences the effect of DAF on timing variability, with stronger effects on bimanual than unimanual coordination, but not on tapping rate. Analyses of movement data suggest that DAF effects on tapping variability were well predicted by the height of the finger that generated the feedback tone (the tap preceding the DAF onset) but not by the finger preparing for the next tap (the tap occurring after the DAF onset). These results suggest that DAF affects movement timing based on the coordination between actions and their effects, rather than by interfering with execution as suggested elsewhere (e.g., Howell, 2004, Contemp. Issues in Comm. Sci. and Disorders).

Email: pqp@buffalo.edu; caroline.palmer@mcgill.ca; simone.dalla-bella@univ-montp1.fr; btkraus@buffalo.edu

10

Perceptual Evaluation of a Binaural Beamforming Algorithm

Jumana Harianawala* *Starkey Hearing Technologies*
Susie Valentine *Starkey Hearing Technologies*
Ivo Merks *Starkey Hearing Technologies*

Improving speech understanding in noise for listeners with hearing impairment has long been a goal of various signal processing algorithms. Such an improvement can be achieved by using a microphone array within each hearing aid (i.e. bilateral directional hearing aids) or by forming a microphone array using two hearing aids (i.e. binaural directional hearing aids). The binaural advantage for speech recognition in reverberant environments and a background of noise is very well documented for both listeners with and without hearing loss (Bronkhorst & Plomp, 1992; Yost, 1997). The magnitude of this binaural advantage (~3 dB) remains unchanged with bilateral omnidirectional or bilateral directional hearing aids (Ricketts, 2000c). Binaural directional hearing aids may be able to provide additional benefits for speech recognition. However, such benefits may disappear if binaural cues such as interaural time difference (ITD) and interaural level difference (ILD) are not properly preserved in the binaural beamforming design. The goal of this study was to investigate the perceptual performance of a binaural directional hearing aid as compared to bilateral hearing aids. The following conditions were assessed for speech intelligibility, localization and spatial release from masking performance: a. Omnidirectional b. Bilateral Directional c. Binaural Directional with partially preserved binaural cues d. Binaural Directional with no preservation of binaural cues The results will be discussed with the implications for binaural beamforming design trade-offs.

Email: jumana_harianawala@starkey.com; susie_valentine@starkey.com; ivo_merks@starkey.com

11

Integration of Vestibular and Auditory Input in the Interpretation of Dynamic Sound Localization Cues

Janet Kim* *Western University*
Ewan Macpherson *University of Western Ontario*

Information about the location of a sound source in the front/back dimension exists in the relation between a listener's head rotation and the resulting changes in interaural time- or level-difference cues. The use of such dynamic localization cues may significantly involve the vestibular system, as the listener needs to maintain an accurate spatial representation of the orientation of their head during head motion. We measured, in both active and passive rotation conditons, the ability of normally hearing human listeners to localize front and back sources of low-frequency (0.5-1 kHz) noise band that typically cannot be accurately localized without head motion. These targets were presented during head rotations with velocities of 25-100 deg/s and were gated on and off as the head moved through a variable-width spatial window ranging between 2.6-40 deg. Performance improved monotonically with increasing spatial window width, which provided larger changes in interaural cues, but decreased as the velocity of head rotation increased, which reduced the stimulus duration. These effects were almost exactly reciprocal in both active and passive rotation conditions, such that performance was primarily related to the stimulus duration, with about 100 ms duration required for 75% correct front/back discrimination regardless of the cue-change magnitude or method of head rotation. The similarity in localization accuracy in the passive rotation condition, which minimized information about rotation carried by efference copy or proprioception, and to the active rotation condition suggests that vestibular input is sufficient to inform the auditory system about head movement.

Email: jkim2223@uwo.ca; ewan.macpherson@nca.uwo.ca

12

Auditory Learning-Related Shifts in Generalization: A Case for Distortions of Representation

Matthew G. Wisniewski* *University at Buffalo, SUNY*
Barbara A. Church *University at Buffalo, SUNY*
Eduardo Mercado III *University at Buffalo, SUNY*

Learning to discriminate similar sensory experiences can alter the ability to identify those original experiences from novel ones. In some cases, learning causes shifts in the identification of original stimuli to novel ones that accent critical properties of the originals. Traditionally, these learning-related shift effects are explained with representation-to-decision associations or relational learning processes. In the current experiment, undergraduate students learned to identify one rate of frequency modulation as a "Target" and another slightly slower rate as a non-target sound. They were then played several modulation rates and asked to identify the target rate heard in training. Consistent with the predictions of standard learning theories, the results show a shift in identification to faster modulation rates. However, an analogous shift was found using a three-alternative forced-choice task in which previously learned response-to-decision associations and relational learning strategies cannot be applied. This suggests that mechanisms contributing to learning-related shifts may be representational. That is, the learning that leads to shift may involve changes to sensory representations, rather than (or in addition to) how decisions based on those representations are made.

Email: mgw@buffalo.edu; bchurch@buffalo.edu; emiii@buffalo.edu

13

Automatic Classification of Multi-Class Acoustic Environments

Martin F. McKinney* *Starkey Hearing Technologies*
Anil Shankar *Starkey Hearing Technologies*
Tao Zhang *Starkey Hearing Technologies*

Hearing aid wearers benefit from technology that can automatically detect and adapt to the surrounding acoustic environment. Typical methods for computationally-efficient environment classification often parcel the acoustic environment into single holistic classes and fail when encountering complex environments with multiple sound sources. We address this deficit in two ways: 1) we examine novel methods for constructing a robust training database of sound clips from environments with multiple sound sources; 2) we examine the ability of different classification architectures to deal with simultaneous sound classes. In the development of classification systems, the integrity of the training database is of utmost importance. With complex acoustic scenes, it becomes difficult to accurately label the relative presence or importance of various sound sources. We used crowd-sourcing methods to label a training database acoustic scenes and evaluate different methods for eliciting the data. Results show that subtle differences in methods can lead to varying results and impact the interpretation of the data. Finally, we examine different classification architectures in their ability to deal with multiple simultaneous classes. Results show that the basic form of the classifier can contribute to the performance in complex acoustic settings.

Email: martin_mckinney@starkey.com; anil_shankar@starkey.com; tao_zhang@starkey.com

14

Auditory Stimulus Generation Tools in MaxforLive

Michael D. Hall* *James Madison University*
Christopher Becker *James Madison University*
Thomas Redpath *James Madison University*
Ashley Assgari *James Madison University*

At APCAM (Hall, Redpath & Becker, 2011), we recently introduced a parallel formant-based, subtractive synthesizer to address needs for rapid stimulus generation in psychoacoustic research. This device avoids limitations of speech synthesizers (e.g., cascaded filters and articulatory parameters), as well as commercial musical synthesizers (e.g., difficult to use or lacking filters needed to adequately model natural sources). It was programmed as a MIDI-controllable, monophonic instrument in MaxforLive, which brings Max/MSP to Ableton's sequencing software Live. The synthesizer included an accompanying additive device, SourceBuilder, that could generate 100-harmonic tones (with assignable frequencies, amplitudes, and phases) that then could be imported as oscillators within Formant Function. Over the past year, the synthesizer has been extensively improved, including a re-designed interface and sub-patchers to simplify the understanding of signal routing. Other modifications include sinewave oscillators, independent control of amplitude and spectral envelopes, a state-variable filter, assignable L/R delays, panning, and chorusing effects, as well as glide time for changes in fundamental frequency, typing names of presets, and noise-free preset switching. Several other devices have been developed to permit synthesis of particular stimuli, as well as more detailed modeling of natural sources. For example, Formant Flux permits text input (which can be pasted from spectral analyses) of formant center frequencies and bandwidths, as well as amplitude information, with linear interpolation between entered values to within a millisecond. A Gunshot Simulator was made possible by the inclusion of rapidly gated filtering. In another device (currently entitled Bottlenose) the frequency range was extended, and a string of fundamental frequency-time pair values could be accepted to permit a reasonable approximation of dolphin calls. These devices can be ported to both stand-alone and open-source alternatives. Most are freely available from, and will be supported by, the first author, and there are plans for continued expansion of their capabilities.

Email: hallmd@jmu.edu; chris.becker4@gmail.com; caseyjonestheengineer@gmail.com; assgaraa@dukes.jmu.edu

15

Melodic Continuation In Three Dimensions: Comparing Expectation in Sequences of Pitch, Brightness, and Loudness

Jackson E. Graves* *University of Minnesota*
Christophe Micheyl *University of Minnesota*
Andrew J. Oxenham *University of Minnesota*

On hearing a sequence of pitches, listeners develop expectations for how that sequence will continue. Research on melodic continuation generally supposes two kinds of factors: the top-down influence of perceived tonality and the bottom-up influence of melodic contour (relative size and direction of the intervals). For bottom-up, contour-based factors, there is converging evidence that melodies with good continuation tend to have small intervals between notes and narrow overall ranges. Since melodic contour can also be perceived in sequences of notes varying in brightness (an aspect of timbre or sound quality) or loudness instead of pitch, it is reasonable to suppose that the same contour-based expectations that apply to pitch sequences also apply to brightness and loudness sequences. The present study found that subjective continuation ratings for brightness and loudness sequences generally conform to the same contour-based expectations as pitch sequences: preferences for narrow ranges and small intervals emerged in all three dimensions. This outcome is compatible with the hypothesis that perception of melodic contour is a general auditory phenomenon that is not unique to pitch.

Email: grave276@umn.edu; cmicheyl@umn.edu; oxenham@umn.edu

16

Speaker Recognition with Pitch-Shifted Voices

Elan Barenholtz* *Florida Atlantic University*
Brittany Guidone *Florida Atlantic University*

In the current study, we investigated the effect of global pitch shifting on voice identity matching of unfamiliar speakers. On each trial, participants heard recordings of two spoken sentences in sequences. On half of trials, the two sentences were recorded from the same (unfamiliar) individual speaker and on the other half they were recorded from the two different speakers. Participants had to determine whether the two sentences were recorded from the same person or not. We compared performance in the matching task when the voice recordings had been pitch shifted between 0 and 8 semitones (positive or negative; tempo was not altered). Between 0 to 2 semitones, performance in the task was near ceiling but dropped sharply at 3 semitones and remained depressed until a small rise at 6 semitones. These results suggest that voices that have undergone pitch-shifting are perceptually more similar than unaltered voices, making the matching task more difficult. These results may point to an expertise effect for voices falling within standard pitch frequencies.

Email: elan.barenholtz@fau.edu; brittanyguidone@yahoo.com

17

Interactions of Pitch and Timbre: How Changes in One Dimension Affect Discrimination of the Other

Emily Allen* *University of Minnesota*
Andrew Oxenham *University of Minnesota*

Variations in spectral shape, perceived as timbre changes, can lead to poorer fundamental frequency (F0) or pitch discrimination. Less is known about the effects of F0 variations on the discrimination of spectral shape. The two aims of the study were 1) to determine whether the interactions between pitch and timbre were symmetric, and 2) to assess the effects of musical training on listeners' ability to ignore variations in irrelevant perceptual dimensions. Subjects were divided into groups of musicians and non-musicians. In Experiment 1, the subjects' task was to compare sequentially presented tone pairs that differed in either pitch or timbre and to judge which was higher. The difference limens (DLs) obtained for both tasks were used in subsequent experiments. In Experiment 2, F0DLs were measured as a function of the size of random variations in spectral centroid, and vice versa. In Experiment 3, sensitivity was measured as the target parameter and the interfering parameter varied by the same amount, in terms of individual DLs. Results showed that both pitch and timbre DLs were strongly affected by random variations in the non-target dimension. The amount of interference observed was similar for both pitch and timbre dimensions. Although musicians had lower (better) F0DLs than non-musicians on average, the amount of interference produced by random spectral variations was similar for the two groups. In addition, there was no significant difference in spectral centroid (timbre) DLs between musicians and non-musicians with or without random variations in pitch. Overall, performance was better when the random non-target variation was in the same direction as the target variation (e.g., when an upward movement in timbre was paired with an upward movement in pitch). The results confirm that pitch and timbre are not easily separable as perceptual dimensions of hearing, and that directional changes can be confused across the two dimensions. [Supported by NIH grant R01 DC 05216.]

Email: prac0010@umn.edu; oxenham@umn.edu

18

Temporal Preparation and the Perceived Duration of Auditory Oddballs

Elisa Kim* *Michigan State University*
Katherine B. Jones *Michigan State University*
J. Devin McAuley *Michigan State University*

When an unexpected (oddball) stimulus is presented within a series of otherwise identical (standard) stimuli, the duration of the oddball tends to be overestimated. Past studies have attributed oddball overestimation to increased attention to the oddball or habituation to the repeated standard. However, the role of temporal preparation in perceived oddball duration has not been considered. Previous work indicates that temporal preparation enhances attention at expected time points. Several factors affect temporal preparation, including foreperiod and probability of stimulus occurrence. Previous work in our lab suggests that with a variable foreperiod, later-position (longer-foreperiod) oddballs are responded to faster and overestimated to a greater degree than earlier-position (shorter-foreperiod) oddballs. In the experiments presented here, we examined the influence of temporal preparation on the oddball effect by varying oddball position (foreperiod) and the probability that an oddball occurred on each trial. Participants heard a series of 9 tones with a fixed 700-ms inter-onset interval. The oddball, identified by a different pitch, occurred in the 5th-8th position. In contrast to past oddball effect studies, catch trials with no oddball were included, lowering the probability of oddball occurrence. Proportion of catch trials was manipulated as a between-subjects variable. In Experiment 1, participants responded to the oddball as quickly as possible. In Experiment 2, participants judged whether oddball duration was shorter or longer than the standard. We predicted that participants would respond faster to later-position oddballs and perceive them to last longer than earlier-position oddballs, but that catch trial presence would mitigate the effect of position, and that the effect of position would decrease with increasing proportion of catch trials. Preliminary results suggest that the effect of position is indeed diminished by catch trial inclusion. This result supports the view that temporal preparation influences detection and perceived duration of auditory stimuli.

Email: kimelis1@msu.edu; jones.katherine.b@gmail.com; dmcauley@msu.edu

19

Neural Activity Relates to Melodic Interval Probability.

Roger E. Dumas* *University of Minnesota*
Arthur C. Leuthold *Brain Sciences Center, Minnesota Veterans Health Care System*
Joshua K. Lynch *University of Minnesota*
Apostolos P. Georgopoulos *University of Minnesota*

The magnetoencephalography (MEG) signal of listeners to tonal sequences was used as the dependent variable in an auto-regression analysis of melodic interval probability (MIP) as calculated from a corpus of 10,000 musical themes. The analysis yielded highly significant relations between the MEG signal and the MIP. Subjects. Each of the ten healthy, right-handed, non-musician subjects (25-37, F=3) listened to a tone sequence comprising a random permutation of 240 pure tones (0.25 s each) selected from a set of pitches in the key of C major (or A natural minor, 2-octave range: freq. 261.6 Hz - 987.86 Hz, note names C4 - B5, MIDI #60-83). We used a customized MATLAB program to generate ten 10-sequence sets. Each set comprised ten sequences having serial correlations ranging from 0.0 to 0.9 over 5 lags, and each subject heard a unique, randomly-ordered set. Data Collection. Magnetic field activity of the brain was recorded using 248 axial gradiometers (Magnes 3600WH, 4-D Neuroimaging, San Diego, CA) at a sampling rate of 1017 Hz (bandpass: 0.1-400 Hz). Given the 666 s stimulus and the sampling rate above, the data generated for each subject consisted of 248 matrices (one per sensor) containing ~677,488 measurements each. Melodic Interval Probability matrix. Using the 'Dictionary of Musical Themes' by Barlow and Morgenstern, we created a Melodic Interval Probability matrix. Using MATLAB©, we converted the nearly 10,000 melodies, all in the key of C, to MIDI note number format. We calculated the frequency of occurrence within the entire corpus for every melodic interval (e. g. unison, minor 2nd, major 3rd, etc.) To calculate melodic interval probabilities (MIP), we divided each interval-frequency value by the total number of intervals in the corpus. The resulting 37x37 matrix showed (predictably) higher frequencies for those intervals between key-member notes, i.e. C, D, E, F, G, A and B. Auto-regression analysis. With respect to the neural signal, successive absolute differences in the MEG signal (DMEG) were calculated between successive note means. With respect to the notes, melodic interval probability was calculated between successive notes. We wanted to assess the relation of changes in neural activity associated with each musical interval in the stimulus to the calculated probability of that interval occurring (MIP). For that purpose, we carried out an autoregression analysis (to account for possibly correlated errors), where the dependent variable was DMEG and the independent variable was the -log of the MIP. Results. We found highly significant, positive relations ($p < .0001$) between the neural activity for each melodic interval and the -log of the probability of that interval's occurrence in the corpus. Conclusion. The human brain learns to expect particular musical notes to follow each other in a composed melody. The more 'unexpected' the following note, the greater is the increase in neural activity.

Email: dumas@umn.edu; leuth004@umn.edu; lynch197@umn.edu; omega@umn.edu