Session 5aAA

Architectural Acoustics and Noise: Potpourri

Norman H. Philipp, Chair
Geiler and Associates, LLC, 1840 E, 153rd Cir., Olathe, KS 66062

Contributed Papers

8:00

5aAA1. Defusing the controversy of scattering versus diffusion coefficients. Hendrik D. Gideonse (Sound Recording Technology, U. Mass Lowell, 294A Boston Avenue, Medford, MA 02155, hendrikxix@xix-acoustics.com)

Several methodologies for the testing and analysis of diffusers have been developed including the ISO Scattering Coefficient and the AES Diffusion Coefficient. These coefficients are the source of some controversy today and this paper makes the attempt to investigate the benefits and weaknesses of these tools by using them to research and test a new diffuser shape. Several issues are exposed in using the coefficients as both qualitative and quantitative metrics. The most important of these being problems with the validity of the comparison of the diffuser's behavior to that of a like-sized flat panel. Additionally there appears to be an intuitive disconnect between the perceived diffusive merits shown by polar plots and the numerical value of coefficients derived from the plots.

8:15

5aAA2. Noise mitigation at the Combat Arms Training Facility, Wright Patterson Air Force Base, Dayton, OH. William J. Murphy, Edward L. Zechmann, Chucri A. Kardous (Hearing Loss Prevention Team, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, OH 45226-1998, wjm4@cdc.gov), and Ning Xiang (School of Architecture, Rensselaer Polytechnic Institute, Troy, NY)

The Combat Arms Training Facility (CATF) at Wright Patterson Air Force Base in Dayton Ohio was evaluated for the effect of noise treatment to the interior of the firing range. Measurements were conducted in 2009 and 2010 before and after noise treatment. Reverberant energy in the range was reduced through the installation of cementatious shredded fiber board and basalt rock wool to the walls and ceiling of the range. No modifications were made to the windows and doors connecting the range interior to adjacent rooms. Prior to the application of noise treatment, the reverberation times (RT60) ranged from about 3.5 seconds at 100 Hz to about 1.3 seconds at 10 kHz. Following application of the noise treatments, the RT60 was reduced to about 1.6 seconds at 100 Hz to 0.5 seconds at 10 kHz. The critical distance for speech intelligibility increased from about 12 feet to about 22 feet in the speech frequencies 800 to 4000 Hz. The Articulation loss of consonants was improved from 22.5 for the untreated range to 7.2 for the treated range. The noise treatments reduced the reverberation time, increased the critical distance and improved speech intelligibility in the CATF firing range.

8:30

5aAA3. Architectural acoustics in educational facilities: An empirical study on university classrooms in Egypt. Hadia S. Awad (Architecture, Alexandria University, 131 Ahmed Shawky St., Sidi Gaber, Alexandria 21529, Egypt, hadyaawad@pua.edu.eg), Hania H. Farag (Electrical Engineering, Alexandria University, Alexandria, Alexandria, Egypt), Dina S. Taha, and Mohamed A. Hanafi (Architecture, Alexandria University, Alexandria, Alexandria, Egypt)

There is a lack of acoustic performance standards for educational facilities in Egypt, especially university classrooms, resulting in a very bad acoustical quality. There is also a lack in the scientific research in that field of study in Egypt, although this type of research has been held in many other countries. This paper aims at studying the acoustic performance standards in different countries such as (ANSI-S12.60, 2010, BB93, 2003), and then using them to evaluate the acoustical quality in classrooms in Alexandria, including public and private, new and old constructions. The evaluation was done by three methods; B&K2250 SLM, Computer modeling using Odeon acoustic prediction software, and mathematical calculations using Sabine and Eyring formulae. The empirical study results were then compared to the recommended values by the previously mentioned standards. The study focused on sound levels, noise levels, reverberation time and Signal-to-Noise ratios. The results prove that there is a lack of acoustic design in Egypt. Most of the studied classrooms did not meet the recommended values. The main problem is using inappropriate finishing materials, which provide very high reverberation. Some simple and applicable proposals were suggested to improve the acoustical quality of classrooms. Predictions have shown a significant improvement.

8:45

5aAA4. Measurements of the acoustical properties of iron slag panels as porous media using scale models. Ho Jun Kim, Hyung Suk Jang, and Jin Yong Jeon (Architectural Engineering, Hanyang University, Seoul, Seongdong-gu 133791, Republic of Korea, jyjeon@hanyang.ac.kr)

Iron slag, by-product of smelting iron ore, has been widely recycled to road pavement material and concrete for environmental-friendly effect. In the present study, the acoustical characteristics of iron slag panels (porous media) were investigated by using scale model approach. The scale models of the iron panels were constructed by considering different sizes of slags. The measurements of absorption and scattering/diffusion coefficients using scale models were conducted in a 1:10 scale reverberation chamber. Through the measurements, the effects of diameters and thicknesses of slags on acoustical properties of iron slag panels were investigated.

9:00

5aAA5. Sound power measurements in non-ideal enclosures using acoustic energy density. Daniel R. Marquez, Scott D. Sommerfeldt, Timothy W. Leishman (Physics and Astronomy, Brigham Young University, Provo, UT 84604, marquezdanny77@gmail.com), and Jonathan Blotter (Mechanical Engineering, Brigham Young University, Provo, UT 84604)

Sound power measurements are generally made in reverberation or anechoic chambers using acoustic pressure measurements as outlined in specific ISO or other standards. Reverberation chambers are used to approximate diffuse acoustic fields wherein the sound power is directly proportional to the spatially averaged squared pressure. Anechoic chambers are utilized to create a direct field condition, wherein sound power can also be determined from sound pressure levels located on a measurement surface which envelopes the source. This paper will introduce a method that utilizes acoustic energy density to estimate the sound power produced in non-ideal enclosures when both direct and reverberant energies contribute significantly to the total acoustic field. Since the acoustic energy density in an enclosure is more spatially uniform than the acoustic pressure, this method can achieve the same accuracy in determining sound power with fewer measurement positions when spatially averaging. The results from numerical models of several rectangular rooms of varying acoustical properties will be presented and the accuracy of the method will be addressed

9:15

5aAA6. A comparison of the precedence effect with specular and diffusive reflections. Michael T. Pastore (Architectural Acoustics, Rensselaer Polytechnic Institute, 4 Irving Place, Troy, NY 12180, m.torben.pastore@gmail.com)

Despite early reflections arriving along many trajectories, humans can localize sounds based on the direction of the direct sound source. The 'precedence effect' describes a set of phenomena thought to be involved in this human ability to localize sound in what could be confusing reverberant environments. Acousticians often apply diffusive surfaces to remove echoes and reduce spatial variation in rooms designed for speech and music, yet the perceptual effects of these treatments are not well understood. The ability of listeners to localize sound is crucial to the success of an enclosed acoustic environment. Therefore the precedence effect under diffusive conditions bears some exploration. A psychoacoustic experiment attempts to characterize the effect of diffusion on the precedence effect. Using an acoustic pointer, lead/lag stimuli at inter-stimulus intervals ranging from 0–4 ms are used for comparing listeners' perceived lateralization of a target stimulus in the presence of a single simulated specular or diffuse reflection. Bandpass noise bursts centered at 500 Hz, (ITD $\pm 300\ \mu s$) are used for the creation of all stimuli. Listeners' performance is evaluated using existing cross-correlation models.

9:30

5aAA7. The effect of building reflections on the equivalent noise level (Leq) from traffic on Lake Shore Drive. Eric W. McGowan (Audio Arts & Acoustics, Columbia College Chicago, Chicago, IL 60614, eric.mcgowan@loop.colum.edu)

The objective of this study was to determine if the Leq resulting from vehicular traffic on Lake Shore Drive in Chicago was affected by reflections from buildings being present in the background. Three test locations were chosen along the same stretch of road, two with buildings in the background, one without. A noise propagation model using CadnaA was constructed in order to compare results from the test locations. Traffic noise spectra were recorded at each location for one (1) hour and post-processed using SpectraPLUS; test data were also acquired using a Type II SPL meter with Leq capabilities. The location with the most building coverage in the background yielded the highest Leq, while the location with no buildings gave the lowest level. The computer model produced results that followed the trends of the test, but the predicted values were consistently higher than the measured levels. The study analyzed if the effect of parameters such as pavement and country-specific Standards could account for the differences between measured and modeled data.

9:45

5aAA8. Abstract withdrawn

10:00–10:15 Break

10:15

5aAA9. Computer simulation of Benjamin Franklin's acoustic experiment on George Whitefield's oratory. Braxton B. Boren and Agnieszka Roginska (Music, New York University, New York, NY 10012, bbb259@nyu.edu)

The Anglican preacher George Whitefield was renowned for his loud voice and the huge crowds he drew during the transatlantic revivals of the 18th century. Benjamin Franklin was skeptical of the accounts of crowds of 30,000 gathering in London, and when Whitefield came to Philadelphia in 1739, Franklin performed one of the earliest recorded 'archeoacoustic' experiments: walking backwards down Market Street, Franklin continued listening to Whitefield speak from the old courthouse until his sermon became unintelligible. Using this maximum intelligible distance, Franklin calculated that Whitefield probably could have been heard by more than 30,000 listeners. Using Franklin's account and period maps and prints of the colonial city, we have built a virtual CAD model of Philadelphia as it would have existed during Whitefield's visit. This paper discusses techniques employed using geometric acoustic simulation software to approximate the loudness of Whitefield's voice based on the STI at Franklin's position. To determine the STI, the background noise at Franklin's position is simulated according to his account of a noise source on Front Street. Given a specific noise source and a minimum intelligible STI, this system yields a loudness value in dB-SPL for an acoustic source at Whitefield's position.

10:30

5aAA10. Streamlining sound power level measurements for determining the product noise rating for consumer products. Matthew A. Nobile (IBM Acoustics Lab, M/S P226, Bldg 704, Boardman Road Site, 2455 South Road, Poughkeepsie, NY 12601, nobile@us.ibm.com)

Under the assumption that consumer demand for quieter products will increase once noise level information is routinely available, a simplified product noise rating method has been proposed and described in several recent papers and forums (e.g., Inter-Noise 2011 and Inter-Noise 2012 papers by the author). This "PNR method" will provide useful noise information to the general public to help them make informed purchasing decisions. The essential elements of the new method are: (1) a product noise rating scale, (2) the Product Noise Rating (PNR), itself, (3) a range-of-levels indication, and (4) a visual icon for presenting the PNR value. But no matter how well-defined this new method is for providing noise levels to consumers, it is useless until the levels, themselves, have actually been measured for a wide range of products. Along this line, this paper will suggest the use of streamlined sound power level measurement procedures on typical consumer products that might help speed the acceptance and use of the new PNR method by manufacturers, product resellers, consumer testing organizations, and other stakeholders.

10:45

5aAA11. Acoustic design of transit stations. Alan Oldfield and Frank Babic (AECOM, 5600 Cancross Court, Suite A, Mississauga, ON L5R 3E9, Canada, alan.oldfield@aecom.com)

The acoustical environment in transit stations is being given increasing attention, particularly in national codes and client standards which now include requirements for intelligibility of voice alarm systems. This differs from the past where fewer loudspeakers with greater sound pressure output levels could be used to fulfill overall sound level criteria, with little consideration for alarm intelligibility. Room acoustics in stations is being considered using techniques previously reserved for performance venues. In addition, noise from tunnel ventilation systems for exhaust and emergency smoke extract is a growing concern that requires attention to meet strict criteria in often constrained spatial conditions and demanding environmental requirements. Incorporating acoustic treatment is often given a low priority over considerations for cost, physical security and maintenance, not to mention the architect's aesthetic vision. This paper presents a review of the key challenges and best practices in acoustic design of transit stations.

11:00

5aAA12. Effects of short noise bursts on human performance and perception. Christopher Ainley and Lily M. Wang (Durham School of Architectural Engr. and Constr., University of Nebraska - Lincoln, 241 Peter Kiewit Institute, 1110 S. 67th Street, Omaha, NE 68182, cainley@unomaha.edu)

The goal of this research project is to better quantify human reactions to short bursts of noise, to complement research at NASA Langley Research Center on evaluating human response inside buildings to low-level sonic booms. The project involved exposing participants over 30-minute sessions to 250 ms broadband noise bursts of certain levels, presented in a controlled yet randomized fashion throughout the session, and gathering responses on human perception and performance on an arithmetic task dealing with short-term memory. While previous research has demonstrated effects of noise bursts of varying amplitudes on other types of tasks that study cognitive processing including attention and at louder levels on this arithmetic task (i.e. 100 dB peak), more information is needed to indicate at what level and to what degree such noise bursts may impact human performance and perception. Twenty-seven test subjects were tested over multiple sessions, with four different levels of the noise bursts, ranging from 47 to 77 dBA. The gathered performance and perception results will be presented, with consideration of each subject's self-reported noise sensitivity. The results will also be correlated to a number of noise metrics, including perceived level (PL) and L1-L99. [Work supported by a NASA Nebraska Space Grant.]

11:15

5aAA13. Community and individual variation in response to noise from high amplitude impulsive sounds. Edward T. Nykaza and Dan Valente (US Army Corps of Engineers, Engineer Research and Development Center, 2902 Newmark Drive, Champaign, IL 61822, edward.t.nykaza@erdc.dren.mil)

It is common for residents living on and around military installations to be exposed to a significant amount of high amplitude impulsive noise, primarily from large weaponry and other blast noise producing sources. Yet in comparison to transportation noise, there have been relatively few studies of how communities and individuals respond to this type of noise. This presentation will report the latest findings from recent human response to blast noise studies conducted at three military installations. Across all sites, blast noise has been found to be the most annoying noise source, despite the fact that a large percentage of respondents reported that their neighborhood was a good or excellent place to live. It has also been found that each community and individual has a unique tolerance to blast noise. Furthermore, individuals use a different and finite portion of the response scale, suggesting that the current methodology of fixating on the percent of the population that is highly annoyed may inadvertently be discarding useful response information. Comparisons between respondents living on- and off-post within and between study sites will be made, with special emphasis placed on differences between the community tolerance level and the community tolerance spread for each site.

11:30

5aAA14. Whoosh noise measurements from an automotive centrifugal compressor. Rick Dehner, Neil Figurella, Kevin Fogarty, and Ahmet Selamet (Mechanical Engineering, The Ohio State University, Columbus, OH 43212, dehner.10@osu.edu)

Broadband noise, accompanying the flow separation and turbulence, is studied for an automotive centrifugal compressor on a steady-flow turbocharger bench facility. When operated in the mid-flow region, the current compressor exhibited elevated broadband noise at high frequencies, which was evident both in the upstream compressor duct and external sound pressure level (SPL) measurement locations. Viewing SPL data as a function of the flow angle relative to the leading edge of the inducer blades (incidence angle) reveals a relationship which is nearly independent of the rotational speed. As the incidence angle is increased, broadband noise levels first go up gradually, nearly level off, and then decrease sharply at a critical incidence angle.

11:45

5aAA15. A study of background noise variability in acoustical measurement laboratories. Seth Bard (IBM Hudson Valley Acoustics Lab, M/S P226, Bldg 704, Boardman Road Site, 2455 South Road, Poughkeepsie, NY 12601, sethbard@us.ibm.com)

Presently the background noise level is included the background correction procedures in the noise emissions standards, ISO 3741, ISO 3744, ISO 3745, and ISO 11201, but the variability of the background noise is not assessed in the corrections. This paper continues the work presented in a previous paper [Proc. Inter-Noise 2012, Paper 1045] by identifying the standard deviations of the background noise sound pressure levels in multiple acoustical test laboratories for the purpose of defining the measurement capability. The foundational paper proposes using both the variability of the background noise and the difference between the source-plus-background measurement and the background measurement to calculate the background noise correction. The data obtained in the background noise study are converted into sound pressure level minima for a source measurement conforming to the requirements of the new proposal in the aforementioned foundational paper. These results are then compared with the smallest source measurements conforming to the current noise emission standards.

Session 5aSC

Speech Communication: Speech Production II: Anatomy, Models, and Methods (Poster Session)

Jody E. Kreiman, Chair
UCLA, Los Angeles, CA 90403

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

5aSC1. Progression of age-related voice instabilities in a non-pathological voice. Eric J. Hunter and Ingo R. Titze (National Center for Voice and Speech, University of Utah, Salt Lake City, UT 84101, eric.hunter@utah.edu)

Age-related structural and functional changes to the aerodigestive tract can affect breathing, swallowing, and voice. Not only can these changes can shape an individual's quality of life, they can, ultimately, be life-threatening. Looking at the voice specifically, changes to the subglottal and supraglottal airways influence vocal fold vibration by producing major bifurcations in the vibration regime. These bifurcations are evidenced in the voice by aphonic segments, subharmonic or side-band frequencies, frequency jumps, and chaotic vibration. The increased occurrence of these bifurcations can, in turn, may indicate age related changes in the vocal folds. The current study examines age-related changes in voice production in two individuals spanning 48–98 y/o and 52–90 y/o. Previous studies revealed changes in breath rate and pitch begin between the ages of 68–74 y/o, indicating a fundamental change in the body's maintenance of the speech mechanism. Voice breaks and bobbles are shown to correspond with this change. These voice breaks, along with the increased voice pitch previously reported, may indicate an interaction of the subglottal and supraglottal airway and an increased weakness of the vocal folds. Weakened vocal folds could also indicate compromised swallowing and airway protection mechanisms.

5aSC2. Spectrographic analysis as an indicator of perceived progress during speech therapy. Kathleen Siren (Speech-Language Pathology/Audiology, Loyola University Maryland, 4501 North Charles St., Baltimore, MD 21210, ksiren@loyola.edu) and Shannon Katz (Loyola Univ., Baltimore, MD)

Despite a respectable data set in speech literature documenting acoustic features of disordered speech, there is a notable scarcity of investigations documenting use of acoustic analysis during speech therapy. In 1999, Kent, Weismer, Kent, Vorperian and Duffy predicted that "joint perceptual-acoustic analysis" would soon be common in clinics due to the lower cost and ease of use of acoustic analysis procedures. Given the paucity of published data, such joint analyses are either still not occurring or are limited to use in clinical settings where information is not shared. Although some clinicians are beginning to use spectrograms as visual biofeedback during speech therapy for children, few studies have documented the effectiveness of this clinical use. Additionally, no studies have attempted to assess the relationship between spectrographic change and change in clinician perception of sound production. This investigation will assess the utility of spectrographic analysis in a clinical setting with children as they progress through speech therapy for misarticulation of /s/ by measuring the acoustic changes occurring over time as sound production matures, and providing additional information about the relationship between acoustic features and perception of accurate sound production.

5aSC3. Variance in tongue motion patterns during the production of /s/. Cindy Ding, Jonghye Woo (Neural and Pain Science, University of Maryland School of Dentistry, Baltimore, MD 21201), Hegang Chen (Epidemiology, University of Maryland School of Medicine, Baltimore, MD), and Maureen Stone (Neural and Pain Science, University of Maryland School of Dentistry, 650 W. Baltimore St Rm 8207, Baltimore, MD 21201, mstone@umaryland.edu)

One of the issues in speech motor control is the nature of variance that occurs in the production of a single speech task spoken by multiple speakers. Do they all essentially use the same gesture modified by fine tuning to adjust to their own anatomy, dialect, etc, or are there quite different, motor equivalent, ways to produce the same speech sound. Production of /s/ in American English is known to be produced using two methods: apical or laminal. Apical /s/ primarily elevates the tongue tip, while laminal /s/ utilizes the tip and blade. Both gestures are found frequently in normal speakers. The present study uses principal components analysis of midsagittal velocity fields to identify the patterns of variance in the internal tongue motion patterns of 10 normal speakers. Palate height will also be examined, as preliminary evidence points to low-palate speakers having a preference for apical /s/, while high-palate speakers use either. Tagged-MRI was used to record 'a geese', and the motion between /g/ and /s/ studied for amount and directions of variance. The goal is to identify stable features of the motion and the effects of /s/ type and palate height.

5aSC4. Single motor unit activity in the genioglossus muscle during vowel articulation. Amy LaCross, Sayoko Takano, Ian J. Kidder (Department of Physiology, University of Arizona, Tucson, AZ 85701), Peter J. Watson (Speech, Language, Hearing Sciences, University of Minnesota, Minneapolis, MN), and Elizabeth Fiona Bailey (Department of Physiology, University of Arizona, 1501 N. Campbell, Rm. 4104, PO Box 245051, Tucson, AZ 85701, ebailey@email.arizona.edu)

There is an abundance of previous electromyographic (EMG) research conducted in the human tongue that has as its focus the extrinsic tongue muscle, genioglossus (GG) and its role in preserving upper airway patency for purposes of gas exchange. Comparatively few studies have documented GG EMG activities in the performance of volitional tasks such as speech production (Honda 1992; Honda et al. 1992; Kusakawa et al 1993). Here we report on de novo efforts to characterize GG single motor unit (SMU) activities in the context of volitional i.e., speech and automatic i.e., central pattern generator driven respiratory activities. Using tungsten microelectrodes inserted into the belly of the GG in 5 healthy adults, we recorded SMU activity during static articulation of the vowels [a] and [ae] in three conditions: phonated, whispered, and articulation using an electrolarynx. Preliminary findings provide evidence of recruitment and de-recruitment of GG motor units coincident with the onset and offset of vowel articulation in phonated and whispered, but not electrolarynx conditions. Furthermore, fluctuations in GG MU firing rates mirror intensity variations within phonated and whispered utterances. These findings provide much-needed new insights into the differential modulation of lingual motor unit activities for purposes of speech production versus breathing.

5aSC5. Numerical investigation of the influence of thyroarytenoid and cricothyroid muscle contraction on the geometry and biomechanical properties of the vocal folds. Jun Yin and Zhaoyan Zhang (UCLA School of Medicine, 1000 Veteran Ave, 31-24 Rehab Center, Los Angeles, CA 90095, zyzhang@ucla.edu)

It is generally accepted that different voice types can be produced for different stiffness and geometry conditions of the vocal folds. However, it remains unclear how the stiffness and geometry of the vocal folds are regulated through laryngeal muscle activation during phonation. A better understanding of such muscular mechanisms of regulating vocal fold properties would provide important insights into the process of physiological control of phonation. In this study, the influence of the activation of the thyroarytenoid (TA) and cricothyroid (CT) muscles on vocal fold geometry and stress distribution within the vocal fold was investigated in a three-dimensional body-cover continuum model of the vocal folds. The results showed that different combinations of the TA and CT activation levels led to different body-cover stress distributions within the vocal fold. Contraction of the TA muscle also caused the vocal fold to bulge towards the glottal midline and created a medial compression force. The results also showed that, in some conditions, coordination of different laryngeal muscles is required to produce an optimal effect on vocal fold geometry or stiffness. [Work supported by NIH.]

5aSC6. On parameterizing glottal area waveforms from high-speed images. Gang Chen (Department of Electrical Engineering, University of California, Los Angeles (UCLA), Los Angeles, CA 90095-1594, gangchen@ee.ucla.edu), Jody Kreiman, Bruce R. Gerratt (Division of Head and Neck Surgery, School of Medicine, University of California, Los Angeles, CA), Yen-Liang Shue (Dolby Australia, McMahons Point, NSW, Australia), and Abeer Alwan (Department of Electrical Engineering, University of California, Los Angeles, CA)

Because voice signals result from vocal fold vibration, perceptually-meaningful vibratory measures should quantify those aspects of vibration that correspond to differences in voice quality. In this study, glottal area waveforms were calculated from high-speed images of the vocal folds. Principal component analysis was applied to these waveforms to investigate the factors that vary with voice quality. Results showed that the first two principal components were significantly ($p < 0.01$) associated with the open quotient and the ratio of alternating-current to direct-current components. However, these conventional source measures, which are based on glottal flow, do not fully characterize observed variations in glottal area pulse shape across different glottal configurations, especially with respect to patterns of glottal closure that may be perceptually important. A source measure, the Source Dynamic Index (SDI), is proposed to characterize glottal area waveform variation for both complete and incomplete glottal closures. Analyses of "glide" phonations in which quality varied continuously from breathy to pressed showed that the SDI is able to characterize the corresponding continuum of glottal area waveform variation, regardless of the presence or absence of glottal gaps. [Work supported in part by NSF and NIH.]

5aSC7. Incorporating cepstral peak prominence as an acoustic method for assessing variation in voice quality. Katherine L. McDonald and Erik R. Thomas (English, North Carolina State University, Raleigh, NC 27695, erthomas@ncsu.edu)

Cepstral Peak Prominence (CPP) has been proposed as a means of assessing breathiness from recordings (Hillenbrand et al., J. Speech and Hearing Res. 37(1994):769-78). CPP involves computing the deviation of the cepstral peak in a smoothed cepstrum from a regression function for the smoothed cepstrum. It has not been tested extensively, though the few evaluative studies, involving comparisons against auditory judgments of voice quality, have yielded positive results. Previous evaluations have not assessed the use of CPP for comparisons of inter-group or intra-individual variation. A reliable way to gauge breathiness is useful not only for therapeutic purposes, but also for informing researchers about a speaker's repertoire and identity construction. This study examines the efficacy of CPP for analyzing social variables at the individual and group levels, using African American and European American subjects producing read and spontaneous speech under controlled conditions. Results show intra-speaker consistency across speaking tasks, but no consistent inter-ethnic differences. These findings suggest that some aspects of phonation may be more important for individual identity than for group identity.

5aSC8. Perceptual evaluation of voicing source models. Jody E. Kreiman, Bruce R. Gerratt (Head and Neck Surgery, University of California, Los Angeles, 31-24 Rehab Center, 1000 Veteran Avenue, Los Angeles, CA 90403, jkreiman@ucla.edu), Gang Chen (Electrical Engineering, University of California, Los Angeles, CA), Mark Garellek (Linguistics, University of California, Los Angeles, CA), and Abeer Alwan (Electrical Engineering, University of California, Los Angeles, CA)

Many models of the glottal source have been proposed, but none has been systematically validated perceptually, so that it is unclear whether deviations from perfect fit have perceptual importance. If model fit fails in ways that have no perceptual significance, such "errors" can be ignored, but poor fit with respect to perceptually-important features has both theoretical and practical importance. To address this issue, we fit 6 different source models to 40 natural voice sources, and then evaluated fit with respect to time-domain landmarks on the source waveforms and details of the harmonic voice source spectrum. We also generated synthetic copies of the voices using each modeled source pulse, with all other synthesizer parameters held constant, and then conducted a visual sort-and-rate task in which listeners assessed the extent of perceived match between the original natural voice samples and each copy. Discussion will focus on the specific strengths and weaknesses of each modeling approach for characterizing differences in vocal quality. [Work supported by NIH/NIDCD grant DC01797 and NSF grant IIS-1018863.]

5aSC9. Improving vocal tract reconstruction and modeling through super-resolution volume reconstruction technique. Jonghye Woo (Departments of Neural and Pain Sciences and Orthodontics, University of Maryland Dental School, Baltimore, MD), Xinhui Zhou (Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20740, zxinhui2001@gmail.com), Maureen Stone (Departments of Neural and Pain Sciences and Orthodontics, University of Maryland Dental School, Baltimore, MD), Jerry L. Prince (Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD), and Carol Y. Espy-Wilson (Department of Electrical and Computer Engineering, University of Maryland, College Park, College Park, MD)

Magnetic resonance imaging has been widely used in speech production for vocal tract reconstruction and modeling. In order to observe detailed structures in the vocal tract, three orthogonal image stacks (sagittal, coronal, and axial) are usually acquired. Due to many constraints, each stack typically has an in-plane resolution which is much better than the out-of-plane resolution. Usually vocal tract modeling is based on just one of these three stacks. As a result, additional useful information revealed by the other two datasets is excluded in the vocal tract model. This study is to improve the vocal tract reconstruction and modeling by integrating information from all of the three stacks. To do so, a super-resolution reconstruction method recently developed to generate an isotropic image volume is used to integrate the three orthogonal stacks. Based on the ATR MRI database of vowel production, vocal tract models from MR images in high resolution, low resolution (simulated through downsampling), and super-resolution were built respectively and compared. The improvement in vocal tract modeling due to the super-resolution technique will be demonstrated on five vowels in terms of visualization and acoustic responses. [This research was supported by NIH R01 CA133015.]

5aSC10. Intraglottal velocity and displacement measurements in an excised larynx. Liran Oren, Ephraim Gutmark (Aerospace Engineering and Engineering Mechanics, University of Cincinnati, PO Box 670528, Cincinnati, OH 45267, orenl@mail.uc.edu), and Sid Khosla (Otolaryngology - Head and Neck Surgery, University of Cincinnati, Cincinnati, OH)

A major assumption of previously published PIV measurements in excised larynges, is that the vortices seen directly above the glottal exit during closing are due to the flow separation vortices (FSV) that formed in the glottis. This assumption needed experimental validation. In addition, the pressures associated with the vortices above the vocal folds may be different than the pressures generated by the intraglottal FSV. Previous studies also relied on images of the glottal opening taken from above of the folds to evaluate the displacement of the folds. This method cannot separate the dynamic of the folds that occurs along their superior and inferior edges. The current study uses a major modification of the PIV techniques previously described to simultaneously measure intraglottal velocity fields and intraglottal geometry. The intraglottal pressure distribution is computed from the velocity measurements by solving the pressure Poisson equation. The results show that strong negative pressure is formed towards the superior edge of the folds. This negative pressure can produce additional force during closing. The displacement of the folds, during closing, is also extracted from the PIV images and it shows that the acceleration of the superior edge is consistently higher than the inferior edge.

5aSC11. Intraglottal pressures related to glottal and laryngeal asymmetries. Ronald Scherer (Communication Sciences and Disorders, Bowling Green State University, 200 Health Center, Bowling Green, OH 43403, ronalds@bgsu.edu)

The air pressures within the glottis during phonation may be highly dependent upon the symmetry of the glottis and other laryngeal structures. Physical laryngeal models M5 and M6 have been used to explore the slanted "oblique" glottis with different included and oblique angles, the change to three dimensions, and the presence of the arytenoid cartilages. Results suggest that (1) the greater the oblique angle, the higher the glottal entrance pressures compared to the symmetric glottis, (2) the larger the intraglottal angle, the more different the pressures are on the two glottal sides, (3) pressures are higher on the convergent side than on the divergent side of the glottis, (4) intraglottal pressure depends on which side of the glottis the flow exits, with the Flow Wall side having lower pressures, (5) for large glottal diameters, all intraglottal pressures may be lower on the Flow Wall side, (6) for special cases of high transglottal pressure, the pressure near glottal exit may be lower than at entrance (divergent glottis), (7) a symmetric but 3-dimensional glottis has consistent but minor pressure changes in the anterior-posterior direction, and (8) the presence of the arytenoid cartilages has minor effects on the intraglottal pressures. [Support from NIH.]

5aSC12. An articulatory and acoustic study of fricative consonants /s/ and /sh/ in normal and post-glossectomy speakers. Xinhui Zhou (Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20740, zxinhui2001@gmail.com), Woo Jonghye, Maureen Stone (Departments of Neural and Pain Sciences and Orthodontics, University of Maryland Dental School, Baltimore, MD), and Carol Y. Espy-Wilson (Department of Electrical and Computer Engineering, University of Maryland, College Park, MD)

Glossectomy is a surgical procedure to remove the cancerous tumor of the tongue. After the glossectomy, the tongue is sutured closed or a flap is inserted to reconstruct the tongue volume. As a result, the properties of the tongue are more or less affected by the surgery. The changes in the tongue properties may also affect the speech production abilities of the post-glossectomy speaker. This study examined the production of the fricative consonants /s/ and /sh/ in normal and post-glossectomy speakers. The data analyzed consisted of audio and magnetic resonance images from dozens of normal and glossectomy speakers. An acoustic analysis showed that the average centers of gravity of /s/ and /sh/ in glossectomy speakers are significantly lower than in normals. This difference may be explained by a more posterior constriction in glossectomees due to the surgery. Examination of the tongue shapes in midsagittal MR images showed that they tend to have more laminal /s/ than apical /s/. 3-D vocal tracts of /s/ and /sh/ were reconstructed for three glossectomy speakers whose /s/ and /sh/ cannot easily be discriminated in listening tests. Details of the 3-D vocal tract shapes, along with their acoustic implications, will be discussed for the glossectomy and normal speakers. [This research was supported by NIH R01 CA133015.]

5aSC13. Acoustic discrimination of Parkinsonian speech using cepstral measures of articulation. Mark D. Skowronski, Rahul Shrivastav (Communicative Sciences and Disorders, Michigan State University, Lansing, MI 48824, markskow@hotmail.com), James Harnsberger (Linguistics, University of Florida, Gainesville, FL), Supraja Anand, and Jay Rosenbek (Speech, Language, and Hearing Sciences, University of Florida, Gainesville, FL)

The effects of idiopathic Parkinson's disease (IPD) on speech include articulatory imprecision. We quantified articulation rate and range acoustically using cepstral coefficients to represent vocal tract settings. Cepstral coefficients were extracted from 10 sentences spoken by 76 talkers, half of which were diagnosed with IPD. Articulation range was estimated for each sentence as the sum across cepstral coefficients of the standard deviation of each coefficient, and articulation rate was estimated using the same procedure, replacing cepstral coefficients with delta coefficients. The mean $\pm$ standard deviation (N = 380 sentences) for the articulation measures of range (7.95 ± 0.50 vs. 6.66 ± 0.53) and rate (5.64 ± 0.56 vs. 4.40 ± 0.46) were significantly different (t-test, $p < 0.001$) for normal vs. IPD speech, respectively. In a leave-one-talker-out classification experiment, range accuracy was 90.1%, rate accuracy was 88.8%, and accuracy was 92.9% using a combined model of articulation range and rate. The strengths of the articulation measures include 1) sensitivity to IPD speech, 2) reliance on cepstral coefficients which have been used for over 30 years to represent speech, 3) no segmentation requirements, 4) low sensitivity to speech material, and 5) effective with only 2 seconds of speech.

5aSC14. Second-formant locus patterns in dysarthric speech. Heejin Kim and Mark Hasegawa-Johnson (Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801, hkim17@illinois.edu)

Second-formant (F2) locus equations represent a linear relationship between F2 measured at the vowel onset following stop release and F2 measured at the vowel midpoint in a consonant-vowel (CV) sequence. Prior research has used the slope and intercept of locus equations as indices to coarticulation degree and the consonant's place of articulation. This presentation addresses coarticulation degree and place of articulation contrasts in dysarthric speech, by comparing locus equation measures for speakers with cerebral palsy and control speakers. Locus equation data are extracted from the Universal Access Speech (Kim et al. 2008). The data consist of CV sequences with labial, alveolar, velar stops produced in the context of various vowels that differ in backness and thus in F2. Results show that for alveolars and labials, slopes are less steep and intercepts are higher in dysarthric speech compared to normal speech, indicating a reduced degree of coarticulation in CV transitions, while for front and back velars, the opposite pattern is observed. In addition, a second-order locus equation analysis shows a reduced separation especially between alveolars and front velars in dysarthric speech. Results will be discussed in relation to the horizontal tongue body positions in CV transitions in dysarthric speech.

5aSC15. Temporal structure in the speech of persons with Dementia of the Alzheimer's Type. Linda Carozza, Pamela Cantor (Communication Sciences & Disorders, St. John's University, Queens, NY), and Fredericka Bell-Berti (Communication Sciences & Disorders, St. John's University, 8000 Utopia Parkway, Queens, NY 11439, bellf@stjohns.edu)

Although cognitive and language processes in dementia have been studied extensively, the question of motor speech degeneration in the course of dementing illness is a relatively unexplored area. The potential for early dissociation of motor functions of language at the level of speech production has not been explored; an interaction between motor speech and language production and perception changes should inform our understanding of the deterioration in dementia. In previous reports on two persons with DAT, we have shown inconsistent final lengthening and effects of syllable-final consonant voicing on vowel duration for one of the two speakers. We recorded one of the speakers again, and his speech was markedly slower. In this report, we expand our analysis to include three additional persons with mild-to-moderate DAT, from whom a series of 4-word phrases containing a target word occurring in phrase-medial or phrase final position was elicited. We will present the results of our analysis of final lengthening, compensatory shortening, and the effects of final consonant voicing on vowel duration.

5aSC16. Articulatory-to-acoustic relations in response to speaking rate modulation in talkers with amyotrophic lateral sclerosis. Antje Mefferd and Stephanie Entz (Wichita State University, Wichita, KS 67260, antje.mefferd@wichita.edu)

The purpose of this study was twofold. One goal was to determine the effects of speaking rate modulation on tongue kinematic and vowel acoustic distinctiveness in talkers with amyotrophic lateral sclerosis (ALS). Another goal was to determine the strength of articulatory-to-acoustic relations in response to speaking rate modulations in talkers with ALS. Six talkers with mild ALS and six healthy controls repeated "See a kite again" at their habitual rate, at a fast rate and a slow rate. The posterior tongue motion was captured simultaneously with the acoustic signal using a 3D electromagnetic articulograph (AG500). To determine kinematic and acoustic distinctiveness maximum tongue excursions and F1/F2 vowel space distance were calculated for the diphthong "ai" in "kite." Preliminary findings showed a greater effect of rate modulation on acoustic distinctiveness than on articulatory distinctiveness for both groups of speakers. The predictability of acoustic distinctiveness based on articulatory distinctiveness varied greatly amongst talkers of both groups. Findings provide empirical evidence of quantal relations between incremental changes of vocal tract configuration and vowel acoustics. Further, findings yield important clinical implications to improve intelligibility and potential explanations for speaking rate declines in talkers with ALS.

5aSC17. Co-registration of articulographic and real-time magnetic resonance imaging data for multimodal analysis of rapid speech. Jangwon Kim (Electrical Engineering, University of Southern California, 3740 McClintock Avenue, Los Angeles, CA 90089, jangwon@usc.edu), Adam Lammert (Computer Science, University of Southern California, Los Angeles, CA), Michael Proctor, and Shrikanth Narayanan (Electrical Engineering, University of Southern California, Los Angeles, CA)

We propose a method for co-registrating speech articulatory/acoustic data from two modalities that provide complementary advantages. Electromagnetic Articulography (EMA) provides high temporal resolution (100 samples/second in WAVE system) and flesh-point tracking, while real-time Magnetic Resonance Imaging, rtMRI, (23 frames/second) offers a complete midsagittal view of the vocal tract, including articulated structures and the articulatory environment. Co-registration was achieved through iterative alignment in the acoustic and articulatory domains. Acoustic signals were aligned temporally using Dynamic Time Warping, while articulatory signals were aligned variously by minimization of mean total error between articulatometry data and estimated corresponding flesh points and by using mutual information derived from articulatory parameters for each sentence. We demonstrate our method on a subset of the TIMIT corpus elicited from a male and a female speaker of American English, and illustrate the benefits of co-registered multi-modal data in the study of liquid and fricative consonant production in rapid speech. [Supported by NIH and NSF grants.]

5aSC18. Discriminating vocal tremor source from amplitude envelope modulations. Kathy M. Carbonell, Brad Story, Rosemary Lester, and Andrew J. Lotto (Speech, Language & Hearing Sciences, University of Arizona, Tucson, AZ 85721, kathyc@email.arizona.edu)

Vocal tremor can have a variety of physiological sources. For example, tremors can result from involuntary oscillation of respiratory muscles (respiratory tremor), or of the muscles responsible for vocal fold adduction (adductory tremor) or lengthening (f0 tremor). While the sources of vocal tremor are distinct, they are notoriously difficult to categorize both perceptually and acoustically. In order to develop acoustic measures that can potentially distinguish sources of tremor, speech samples were synthesized using a kinematic model of the vocal folds attached to a model of the vocal tract and trachea [Titze, JASA, 75, 570-580; Story, 2005, JASA, 117, 3231-3254]. Tremors were created by modulating parameters of the vocal fold model corresponding to the three types mentioned above. The acoustic measures were related to temporal regularities in the amplitude envelope computed across the entire signal and select frequency bands. These measures could reliably categorize the samples by tremor source (as determined from a discriminant function analysis) even when compound tremors (created from more than one source) were included. These results supply initial support for an acoustic based approach to diagnosing tremor source and further evidence for the rich information about talker characteristics present in the temporal structure of the amplitude envelope.

5aSC19. Principal components analysis comparison of normal and glossectomy movement patterns in multiple tasks. Caitlin R. Gallagher, Jonghye Woo (Neural and Pain Science, University of Maryland School of Dentistry, Baltimore, MD), Hegang Chen (Epidemiology, University of Maryland School of Medicine, Baltimore, MD), and Maureen Stone (Neural and Pain Science, University of Maryland School of Dentistry, 650 W. Baltimore St Rm 8207, Baltimore, MD 21201, mstone@umaryland.edu)

Recent estimates suggest that 34,000 people are diagnosed with oral cancer each year. The lateral border of the tongue is one of the most common sites for lingual cancer and surgery resections both muscles and nerves leading to the tongue tip. One sound that is typically impaired is /s/ as it requires precise tongue shape, location and palatal contact, and small errors are acoustically salient. This study uses Principal Components Analysis (PCA) to compare motion patterns of the internal tongue during the word 'geese'. The study will compare 4 subjects: one glossectomy patient and a matched control who produce an apical /s/, and another pair that produces a laminal /s/. A PCA will be run for all subjects and the principal patterns of variance will be determined. These patterns will be used to identify stable features of the motion and variations due to /s/-type and patient vs control. The complexity of each subject's motion pattern will be studied to determine whether the patients have more variability in their motion due to strategies of motor adaptation, or less variability due to scarring and increased rigidity.

5aSC20. The role of respiratory/phonatory training and acoustic analysis in normalizing speaking rate in a case of scanning speech. Emily Q. Wang, Samantha A. Sanders (Department of Communication Disorders and Sciences, Rush University Medical Center, Chicago, IL 60523, emily_wang@rush.edu), and Hanjun Liu (Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guang Dong, China)

In this single case study, we report the role of respiratory/phonatory training in normalizing speaking rate in a patient with significant ataxic dysarthria, specifically, scanning speech, two years post a closed head injury from an MVA. The patient passed bilateral hearing screening, and scored 30/30 on the Mini-Mental State Exam, AQ of 97.2 on the Western Aphasia Battery-Revised and 58/60 on the Boston Naming Test. The initial speech analysis revealed the following: perceptually, the speech was slow with equal and excess stress and significant phonatory-prosodic insufficiency. Acoustically, the syllable duration was increased and equalized with very little variation in fundamental frequency or intensity. The phonatory-prosodic insufficiency was identified as the underlying deficit and targeted, which resulted in significant improvement in both overall speech intelligibility and naturalness of speech in seven weeks. Both self-rating and acoustic analyses were used throughout as means of feedback which contributed to the success of the treatment. The final acoustic analysis indicated normal syllable duration and speech prosody in spontaneous speech. The pre- and post-treatment audio clips will be used to accompany the acoustic analysis. The role of acoustic analysis in treatment of phonatory-prosodic insufficiency and its implication in dysarthria management in general will be discussed.

5aSC21. The effect of portable DAF usage in daily life on the speech intelligibility of dysarthrias. Eiji Shimura (Department of Speech, Language and Hearing Sciences, Niigata University of Health and Welfare, 1398, Shimami-cho, Kita-Ku, Niigata 950-3198, Japan, shimura@nuhw.ac.jp) and Kazuhiko Kakehi (Institute for Advanced Studies in Artificial Intelligence, Chukyo University, Toyota, Aichi, Japan)

Dysarthria is a neurologic motor speech impairment due to a pathological change of nerve and muscle system. Several rehabilitation methods of speaking rate control have been widely used to improve speech intelligibility for mild dysarthria. Nevertheless, speech intelligibility of the dysarthric patients show almost no improvement outside of a clinic. Delayed auditory feedback(-DAF)is one of the speaking rate control methods. Recently a small portable DAF device has been developed, which enables dysarthric patients to use it in their daily life. In this study, wearing a portable DAF, two dysarthric patients conducted a 20-minute practice per day for three months. In the practice, they instructed to prolong vowel length in DAF usage. Intelligibility tests for single-word utterance and for reading aloud a long sentence were conducted before and after the practices. As a result, the single-word intelligibility score increased from 84.0% to 90.6% in case1, and from 68.3% to 89.7% in case2.

Additionally, vowel space in the F1-F2 space in case2 enlarged. Both the performance of long sentence reading and free conversation will be presented in the meeting. These results suggest that a portable DAF effectiveness in daily use. This work was supported by JSPS KAKENHI 10424889.

5aSC22. Automated prosodic labeling using soft computing. Nicholas Bacuez (French and Italian, University of Texas, Austin, TX 78712, nicholasbacuez@utexas.edu)

Automated prosodic modeling/labeling usually relies on complex algorithms. However, as phonological research suggests, human beings do not process intonation in terms of algorithms but in terms of relative oppositions, linguistically encoded in words such as 'higher', 'lower', 'longer', 'shorter', and reflected by the notation: H, L, H*, L% … I have conceived an automated procedure -the 4 layer structure- based on linguistic human reasoning that locates the tones of each individual sentence in a corpus. The original time/f0 information (1st level) is scaled onto a 100 by 100 cartesian plane preserving relative oppositions of time and f0 (2nd level). At the pretonal-anchoring level (3rd level), syllables are divided into frames within which the highest and lowest points are marked as pretones; the contour is turned into a string of pretones (4 pretones per syllable) anchored to their position in the sentence. Finally, a multi-pass procedure (4th level) iteratively locates highest and lowest pretones forming larger movements until the entire sentence has been labeled. The procedure uses minimal logical and mathematical tools to quantify the difference between pretones ('higher' or 'lower') and the span of movements (consecutive 'lower' or 'higher'). It was successfully applied to various corpora of French.

5aSC23. Effects of preceding consonant features on /ð/ assimilation. Christina Schwaller (English: Linguistics, North Carolina State University, Raleigh, NC 27695, cschwaller@gmail.com)

Sociolinguists have occasionally mentioned regressive assimilation of the voiced interdental fricative /ð/ to a preceding consonant as a feature of some dialects. However, little research has been published on the topic. This study examines effects of preceding consonant features on /ð/ assimilation, using interviews with residents of Robeson County, NC. Tokens of phonemic /ð/ were examined in spectrogram form to identify which ones were assimilated. Impressionistic judgment was used in conjunction with the spectrograms. Discontinuities within consonant clusters served as evidence of unassimilated tokens. For manner of articulation, features such as stop bursts and frication noise were identified. When necessary, measurements of F2 and F3 at adjacent vowel transitions and midpoints were compared to determine place of articulation. Place, manner, voicing, and nasality of preceding consonants were then recorded. A logistic regression indicated the significance of each variable. This showed that preceding consonants affect /ð/ assimilation. Preceding alveolars are most likely to result in assimilation, which was shown to be significant. The difference in manner between fricatives and stops is also significant, as /ð/ is more likely to assimilate to a fricative. Nasality is highly significant, with /ð/ more likely to assimilate to a nasal than to an oral consonant.

5aSC24. Locus equations as measures of consonantal variation in Wisconsin English: Revisiting the vowel-to-consonant transition. Michael J. Fox (English, North Carolina State University, Rayleigh, NC 27695)

Extant literature on vowel-to-consonant (VC) locus equations (defined as a regression line fit to F2 transition measurements of one consonant paired with many vowels) suggests that they lack the high linearity of the corresponding consonant-to-vowel ones (CV) [Sussman et al. 1997]. Measurements of the second formant taken at the vocalic midpoint and last glottal pulse yielded four VC locus equations from 31 respondents from west-central Wisconsin for the syllable-final consonants /d, t, g, k/ (n = 124). Previous work with WI English indicates the existence of a differential in locus equations for the voiced vs. voiceless velar stops [Purnell 2008]. Apical stops are included for comparison across place of articulation (i.e. apical to velar) with no differential expected. High levels of linearity were found for VC locus equations for all the consonants examined. Fitting second-order locus equations [Chennoukh, et al. 1995, 1997] to the coefficients revealed a differential between /g/ and /k/, but not /d/ and /t/. Moreover, discriminant analyses yielded higher classification rates for locus equations than for token level data. These results run counter to previous characterizations of VC locus equations and suggest the potential for the use of locus equations as measures of dialect-specific coarticulation.

5aSC25. Sensitivity of acoustic parameters of /s/ in adolescents. Christine H. Shadle (Haskins Laboratories, 300 George St., New Haven, CT 06511, shadle@haskins.yale.edu), Laura L. Koenig (Comm. Sci. & Disorders, Long Island University, Brooklyn, NY), and Jonathan Preston (Comm. Sciences, Southern CT State Univ, New Haven, CT)

Acoustic parameters that are closely related to the source and filter properties of fricatives have been shown in previous work to be useful for comparing fricative spectra, as for instance in analyzing a database of /s/ productions by typical and misarticulating children aged 10 to 15. Those results were used to study the sensitivity of different parameters designed to measure the same underlying property. One parameter, the frequency of the main front-cavity resonance, is used to compute two others that measure the effect of labialized contexts, and the degree of sibilance; therefore, three methods of finding that resonance automatically were compared. Gender differences were found in the main resonance frequency and in measures of noise source growth during the fricative, so the use of two sets of heuristically-defined frequency bands were compared, as were spectral slope vs. sound level differences. The back-cavity resonances at low frequencies tend to be more prominent in typical children's /s/ spectra than adults', and still more so in certain types of misarticulation. Different measures of that prominence were compared. The parameters most sensitive to the difference between typical children's and adult's productions, and typical and misarticulating children's productions of /s/, will be described.

5aSC26. Segmenting American English V+/l/ and V+/r/ sequences: Methodological implications. María Riera and Joaquín Romero (Estudis Anglesos i Alemanys, Universitat Rovira i Virgili, Av. Catalunya, 35, Tarragona 43002, Spain, maria.riera@urv.cat)

This paper presents some methodological implications for the segmentation of final V+/l/ and V+/r/ sequences in American English stressed monosyllables. Between the two segments in the sequences a transitional schwa-like element characterized by variable durational and spectral values as a function of the preceding vowel, the following consonant and speaking rate can be identified. Given the dynamic nature of this element, problems related to boundary placement often arise. A segmentation method based solely on spectrographic observation and auditory corroboration has proven in previous studies to be too subjective to be reliable. A more objective method based on first derivative curve extraction, which provides us with first derivative formant traces that show peaks of formant change given by velocity maxima and minima, is more suitable for our purposes. However, the decision as to whether to choose F1, F2 or F3 traces as reference points for boundary placement, together with the presence of too many peaks in some cases and of not enough peaks in other cases, poses problems to the segmentation procedure and makes it necessary for the subjective method of segmentation based on spectrographic observation and auditory corroboration to come into play. Both methods thus complement each other.

5aSC27. Using partially separable functions to image spatiotemporal aspects of Arabic pharyngealization. Ryan Shosted (Linguistics, University of Illinois at Urbana-Champaign, 707 S Mathews Ave, Urbana, IL 61801, rshosted@illinois.edu), Maojing Fu (Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL), Abbas Benmamoun (Linguistics, University of Illinois at Urbana-Champaign, Urbana, IL), Zhi-Pei Liang (Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL), and Bradley P. Sutton (Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL)

It has been challenging to estimate the temporal domain of pharyngealization in Arabic. Conventional MRI has limited assessment of dynamic pharyngeal shape during speech. In this study, fast spiral sequences, combined with partially separable functions, were used to achieve a relatively high spatiotemporal resolution (2.2 mm $\times$ 2.2 mm $\times$ 8.0 mm, at a frame rate of 86 fps) during dynamic speech imaging of a single midsagittal slice. One male speaker of Levantine Arabic produced pairs of words that differed minimally by one speech sound: pharyngeal fricative /ħ/ or non-pharnyngeal /b/. Each word was produced 23 times. The temporal extent of pharyngeal tissue displacement associated with /ħ/ was investigated. Sounds were segmented with reference to a simultaneous, noise-canceled acoustic recording. Spatiotemporal maps of differential pixel intensity (interpreted as tissue displacement) were generated for each segment preceding the pharyngeal /

non-pharyngeal test segment. Average differential pixel intensity in the pharyngeal area was then sampled during these preceding segments. T-tests revealed significant differences ($p < 0.01$) up to two segments away from the pharyngeal / non-pharyngeal test segment. This technique should permit investigation of spatiotemporal aspects of pharyngealization across different varieties of Arabic, where distance and direction of pharyngealization are said to vary systematically.

5aSC28. An experimental comparison of fundamental frequency tracking algorithms. Hongbing Hu and Stephen Zahorian (Electrical and Computer Engineering, State University of New York at Binghamton, PO Box 6000, Binghamton, NY 13902, hongbingh@gmail.com)

"Yet another Algorithm for Pitch Tracking -YAAPT" was published in a 2010 JASA paper (Zahorian and Hu). Although demonstrated to provide high accuracy and noise robustness for fundamental frequency tracking for both studio quality speech and telephone speech, especially as compared to other well-known algorithms (YIN, Praat, RAPT), YAAPT has not been widely used, possibly due to the difficulty of using it and uncertainty about its effectiveness for difficult conditions. Therefore, more work has been done to improve the algorithm and especially to improve its functionality and ease of use as MATLAB functions. In the present paper, the current version of YAAPT is presented, along with clear documentation for using it, both stand alone and as a function to be called by another program. Experimentally, YAAPT is compared with YIN, Praat, RAPT, and a cepstrum method for studio bandwidth speech and telephone speech for a variety of noise conditions. Experiments are conducted with multiple databases, including American English, British English, and Mandarin Chinese. For most conditions evaluated, YAAPT gives better performance than the other fundamental frequency trackers.

FRIDAY MORNING, 26 OCTOBER 2012 MARY LOU WILLIAMS A/B, 8:30 A.M. TO 11:45 A.M.

Session 5aUW

Underwater Acoustics and Acoustical Oceanography: Boundary Interaction and Inversion

Jorge E. Quijano, Chair
School of Earth and Ocean Sciences, University of Victoria, Victoria, BC V8P 5C2, Canada

Contributed Papers

8:30

5aUW1. High-frequency acoustic backscattering from a sand sediment: Experiments and data/model comparisons. Brian T. Hefner, Anatoliy N. Ivakin, and Darrell R. Jackson (Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, WA 98105, hefner@apl.washington.edu)

In the Spring of 2012, high-frequency backscattering from a sandy sediment was measured in the Gulf of Mexico at the site of the upcoming, ONR-sponsored reverberation experiment. The measurements were made using an array of sources and receivers that collected data from 200 to 500 kHz and that could be rotated such that the incident grazing angles varied from 10 to 50 degrees. This array was used previously to measure scattering from a sand/mud sediment during the Sediment Acoustics Experiment 2004 (SAX04). To support data/model comparisons, the seabed roughness, sediment shell content, sediment sound speed, and sediment attenuation were also measured. For scattering below the critical grazing angle, sediment roughness is found to be the dominant scattering mechanism while above the critical angle, roughness scattering underpredicts the measured scattering strength. To understand the scattering strength at high grazing angles, scattering from shells and shell hash is considered. The measured scattering strengths and environmental properties at the experiment site are also compared to those made during SAX04. (Work supported by the US Office of Naval Research)

8:45

5aUW2. The effect of bottom layering on the acoustic vector field. David R. Dall'Osto and Peter H. Dahl (Mechancial Engineering and the Applied Physics Laboratory, University of Washington, Seattle, WA 98103, dallosto@u.washington.edu)

A signal reflected from a layered sea-bed contains information pertaining to the sediment properties. Typically, a signal intended to probe the sea-bed is designed to have a large bandwidth to allow for time separation of arrivals from the multiple layers. Depending on the geometry, it may impossible to avoid interference of these arrivals. The interference of these multiple arrivals does establish a pattern observable in the vector intensity. Measurements of the vertical complex acoustic intensity of a near-bottom source (~λ from the seafloor) collected off the coast of New Jersey in 2006 demonstrate the effect of a sub-bottom layer and the observable interference pattern between the first bottom reflection and the sub-bottom reflection. The spatial structure of the complex intensity can be used to infer bottom properties, which are in close agreement with a number of experimental studies at this location. The observable in the complex intensity can also be directly measured with a particle motion sensor. Parabolic equation simulations of the experimental site are used to demonstrate both the characteristic of the vector field and the sensitivity of these vector properties to changes in the sediment properties.

9:00

5aUW3. A broadband model for a range of ocean sediments. Nicholas Chotiros and Marcia J. Isakson (Applied Research Laboratories, Univ. of Texas at Austin, PO Box 8029, Austin, TX 8713-8029, chotiros@arlut.utexas.edu)

In the context of the Biot theory of sound propagation in porous media, particularly water-saturated granular media, Yamamoto and Turgut [J. Acoust. Soc. Am. 83(5), 1744–1751, May, 1988] have shown that the pore size distribution can have a profound effect on the frequency dependence of the attenuation of sound. In sandy sediments below the characteristic frequency, the attenuation is predicted to increase as the second power of frequency. In soft sediments, it is found that the rate of increase is closer to the first power. By adjusting the width of the pore size distribution, it is possible to smoothly change from the second power of frequency to the first power, in certain frequency bands. This suggests that pore size distribution may be a critical parameter in the determination of sound attenuation in the seabed. The model predictions are compared to measurements from the Shallow Water 2006 experiment as an illustration. [Work supported by the Office of Naval Research, Ocean Acoustics Program.]

9:15

5aUW4. Measurement and modeling of Scholte waves in shallow water. Gopu R. Potty, James H. Miller (Ocean Engineering, University of Rhode Island, Narragansett Bay Campus, Narragansett, RI 02882, potty@egr.uri.edu), and Marcia Isakson (Applied Research Laboratories, The University of Texas at Austin, Austin, TX)

Shear speeds in semi-consolidated and consolidated shallow water sediments can significantly impact compressional wave attenuation and arrival times of acoustic normal modes. In addition shear properties of sediments are directly related to the strength of the sediments in geotechnical applications. All of these factors emphasize the importance of estimating shear speeds in shallow water sediments. One of the most promising approaches to estimate shear speed is to invert the shear speed profile using the dispersion of interface waves (Scholte waves). The propagation speed and attenuation of the Scholte wave are closely related to shear-wave speed and attenuation over a depth of 1–2 wavelengths into the seabed. Data from the tests conducted in Narragansett Bay and off Block Island in water depths ranging from 10 m to 25 m will be presented. Modeling of interface waves will be carried out using Finite Element Method (FEM) and a wave number integration model (OASES). Sediment properties will be inferred based on the modeling and data-model comparison. [Work supported by Office of Naval Research.]

9:30

5aUW5. Validity of first-order perturbation theory for scattering from one-dimensional and two-dimensional rough surfaces described by power-law spectra. Bryant M. Tran (Applied Research Laboratories, The University of Texas at Austin, 10000 Burnet Rd., Austin, TX 78758, bmtran@utexas.edu), Sumedh Joshi (Center for Applied Mathematics, Cornell Univ., Ithaca, NY), and Marcia J. Isakson (Applied Research Laboratories, The University of Texas at Austin, Austin, TX)

First-order perturbation theory is a widely used model for estimating the backscatter of acoustic waves incident on a rough surface. The validity of perturbation theory for one-dimensional surfaces described by Gaussian spectra is well established. However, little has been done to confirm its range of validity when expanded to two-dimensional surfaces. Furthermore, the range of validity for surfaces described by power-law spectra has not been fully explored. This work seeks to benchmark first-order perturbation theory against a finite element method solution for scattering from one-dimensional and two-dimensional rough pressure-release surfaces described by power-law spectra. The relationship between ranges of validity of 1D and 2D surfaces will be considered. [Work sponsored by the Office of Naval Research, Ocean Acoustics.]

9:45

5aUW6. Broadband synthetic aperture matched field geoacoustic inversion with a single hydrophone. Bien Aik Tan, Caglar Yardim, Peter Gerstoft, and William Hodgkiss (Marine Physical Laboratory, Scripps Institution of Oceanography, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0238, btan@ucsd.edu)

Traditionally matched-field geoacoustic inversion experiments sampled the acoustic field on long arrays and require powerful transmissions in order to reduce parameter uncertainty. However, single-hydrophone based geoacoustic inversion methods exist. Practically, these methods are attractive compared to the ones using long arrays. In a single hydrophone setup, spatial diversity is traded off for frequency diversity; the source is broadband. This paper uses single-hydrophone frequency coherent matched-field inversion and exploits extended-duration coherent transmissions (multiple LFM chirps) to increase signal to noise ratio. As a result, the overall signal becomes Doppler/motion intolerant. But Doppler can be modeled by including source and hydrophone horizontal motions. To correlate well with the measured field across a receiver trajectory and to incorporate a transmission across source trajectory, Doppler in waveguide and normal mode theory are applied. The method is demonstrated with 100-900 Hz LFM SW06 data with low signal to noise ratio.

10:00–10:15 Break

10:15

5aUW7. Range and cross-range resolution from a three-dimensional linearized perturbative inversion scheme. Christopher M. Bender, Megan S. Ballard (Appl. Res. Labs, Univ. of Texas at Austin, 10000 Burnet Rd., Austin, TX 78713, bender.christopher@gmail.com), and Preston S. Wilson (Mech. Eng. Dept. and Appl. Res. Labs, Univ. of Texas at Austin, Austin, TX)

The overall goal of this work is to develop a sparse autonomous observing system to sample the four-dimensional ocean. For this purpose, a perturbative inversion scheme [S.D. Rajan, *et. al*., J. Acoust. Soc. Am., **82**, pp. 998-1017 (1987)] is applied to estimate water-column sound-speed in all three spatial dimensions at a single “snapshot” in time. The input data to the inversion are estimates of modal travel time calculated from measurements from a distributed network of acoustic sources and receivers. Temporal variability is assessed by carrying out repeated inversions for new realizations of the input data. In initial applications of the inversion scheme, out-of-plane propagation effects were ignored and the solution was obtained by assuming straight-line paths in the horizontal plane. The effect of neglecting horizontal refraction on solution accuracy is quantified. The vertical and horizontal resolution of the solution depends on the quantity of input data and on the quantity of sources and receivers, respectively. Thus, for a particular environment, the available data may be insufficient to quantify the given variability, resulting in uncertainty in the solution. This talk explores the effects of environmental variability and spatial resolution on the uncertainty of the solution. [Work supported by ARL:UT IR&D.]

10:30

5aUW8. Reconstructing surface wave profiles from reflected acoustic pulses. Sean Walstead (ECE, UCSD, 9500 Gilman Drive, 0407, La Jolla, CA 92093-0407, swalstead@ucsd.edu) and Grant Deane (SIO, UCSD, La Jolla, CA)

Surface wave shapes are determined by analyzing underwater reflected acoustic signals. The acoustic signals (of nominal frequency 200 kHz) are forward scattered from the underside of surface waves that are generated in a wave tank and scaled to model smooth ocean swell. An inverse processing algorithm is designed and implemented to reconstruct the surface displacement profiles of the waves over one complete period. The inverse processing uses the surface scattered pulses collected at the receiver, an initial wave profile (two are considered), and a broadband forward scattering model based on Kirchhoff’s diffraction formula to iteratively adjust the surface until it is considered optimized or reconstructed. Two physical length scales over which information can be known about the surface are confirmed. An outer length scale, the Fresnel zone surrounding each specular reflection point, is the only region where optimized surfaces resulting from each initial profile converge within a resolution set by the inner length scale, a quarter-wavelength of the acoustic pulse. The statistical confidence of each optimized surface is also highest within a Fresnel zone. Future design considerations are suggested such as an array of receivers that increases the region of surface reconstruction by a factor of 2 to 3.

10:45

5aUW9. A true-depth passive fathometer. Jorge E. Quijano, Stan E. Dosso, and Jan Dettmer (School of Earth and Ocean Sciences, University of Victoria, Victoria, BC V8P 5C2, Canada, jorgeq@uvic.ca)

This paper applies a sequential trans-dimensional (trans-D) Monte Carlo algorithm for geoacoustic inversion to bottom-loss data estimated from wind-driven ambient noise at a drifting vertical array. The approach explored in this work provides range-dependent estimates of geoacoustic parameters and true-depth layering structure of the seabed, together with corresponding uncertainties. The Bayesian inversion is applied to incoherent estimates of seabed bottom loss, computed as the array drifts along a range-dependent track. The method adopts a layered representation of the seabed, where each layer is determined by sound speed, density, attenuation, and thickness. The number of layers is also included as an unknown parameter, which allows data-driven parametrization rather than an arbitrary choice of the parametrization for the seabed model. The trans-D Bayesian inversion

method samples the joint posterior probability density of all model parameters to provide parameter estimates and uncertainties. A particle filter is used to update the estimated geoacoustic parameters from one array position to the next. The sequential inversion approach is demonstrated using data from the Boundary 2003 experiment, and compared to images of the seabed layering structure obtained by an active seismic system.

11:00

5aUW10. Sound speed measurement with direct-sequence spread spectrum signal combining time of flight and phase shift. Shen Zhao, Chunjie Qiao, Yue-ke Wang, and Zhi-gang Huang (Department of Instrumental Science and Technology, Mechatronics and Automation School, National University of Defense Technology (NUDT), Chang Sha, Hu Nan 410073, China, quickjobs@163.com)

Increasing demands for high accuracy and rapid measurement of sound speed have prompted the development of portable sound velocimeter in oceanography. Generally, sound speed can be established by CTD or TD (Time Delayed) method respectively. CTD method suffers from errors of corresponding sensors and application range of the converting equations. TD method, sing-around technique representative, is widely used in portable velocimeter. The accuracy of TD is proportional to receiving SNR and signal length. The duration of tone burst in convenient sing-around is limited by multi-echo, thus only a few valid data can be utilized. To address the problems, DSSS (Direct-Sequence-Spread-Spectrum) signal in continuous form is adopted to extract the TD directly, with which sound speed could be established related to laboratory derived equations, Del Grosso's equation representative. Based on this method, two separate transducers were deployed with constant distance. Cross-correlation between received and transmitted (orthogonal form) signals is computed. Combine Time-of-Flight and Phase-Shift, the TD can be derived by two processes, the peak detecting of cross-correlation envelope and carrier phase estimating. Theoretic analyses and experiments indicate that the accuracy with DSSS relates to the period of PN sequence, the frequency of chip, and the linearity characteristic of electronic circuits and transducers.

11:15

5aUW11. An inverse method for estimating sediment sound speed in the ocean. Tao Lin and Zoi-Heleni Michalopoulou (Department of Mathematical Sciences, New Jersey Institute of Technology, 323 ML King Blvd, Newark, NJ 07102, michalop@njit.edu)

In this work, a new inverse method for estimating sediment sound-speed profiles is investigated. Stickler and Deift derived a trace formula for recovering a seiment sound-speed profile by simply using a reflection coefficient at very low frequencies. The method may be sensitive to noise and also involves computationally cumbersome calculations. In our approach, we first design a linear approximation for the trace formula based on the Born approximation, in order to reduce the computational cost. The stability of the modified inverse algorithm is then tested with synthetic noisy data. Finally, we look into ways for relaxing the limiting assumptions of the approach. [Work supported by ONR and the NSF CSUMS program.]

11:30

5aUW12. Application of smoothing techniques to sequential geoacoustic inversion. Caglar Yardim, Peter Gerstoft, and William S. Hodgkiss (Scripps Institution of Oceanography, 9500 Gilman Dr, La Jolla, CA 92093-0238, cyardim@ucsd.edu)

Sequential Bayesian techniques such as particle filters have been successfully used to track a moving source in an unknown, complex, and evolving ocean environment. These methods treat both the source and the ocean parameters as non-stationary unknown random variables and sequentially estimate the best solution in addition to the uncertainties in the estimates. Particle filters are numerical methods called sequential Monte Carlo techniques that can operate on nonlinear systems with non-Gaussian probability density functions. Particle smoothers are a natural extension to the filters. A smoother is appropriate in applications where all data have already been observed and are readily available. Therefore, both past and future measurements can be exploited. Geoacoustic and source tracking is performed here using two smoother algorithms, the forward backward smoother and the two-filter smoother. The approach is demonstrated on experimental data collected during both the SWellEx-96 and SW06 experiments where the uncertainty in the estimates is reduced.