Session 5aAA

Architectural Acoustics, Noise, and INCE: Case Studies, Applications, and Integration of Architectural Acoustics in Building Information Modeling Three-Dimensional Modeling

Norman H. Philipp, Chair
Univ. of Nebraska-Lincoln, Peter Kiewit Inst., 1110 S. 67th St., Omaha, NE 68182-0791

Chair’s Introduction—7:30

Invited Papers

7:35

5aAA1. Building information management: An acoustics consultant's perspective. Benjamin Markham (Acentech Inc., 33 Moulton St., Cambridge, MA 02138)

Architecture and engineering firms use building information management (BIM) software applications increasingly in the design process; similarly, building owners create BIM models for existing facilities both for internal use and to aid future tenants, facilities managers, and designers in planning for maintenance programs, restoration, or renovation. As for other design and engineering disciplines, these tools provide powerful opportunities to the acoustics community. The primary benefits to architectural acoustics consultants fall into two broad categories: first, in the construction process by providing a means of organizing the specifications for materials and constructions, and second, in the design process, by establishing a framework of data and acoustical attributes that in the future could potentially aid in acoustical calculation and analysis.

7:55

5aAA2. Building information modeling with revit architecture and acoustical coordination. Katherine LePage (GWWO, Inc. Architects, 800 Wyman Park Dr., Ste. 300, Baltimore, MD 21211, klepage@gwwoinc.com)

The adoption of BIM has provided a huge shift within the architectural design community. It provides an opportunity for more thoughtfully designed buildings as well as comprehensive coordination with consultants. This presentation will cover the basics of how programs like revit architecture are being used by the architectural community in general and more specifically for acoustical coordination. Topics covered will include visualization, basic modeling, consultant coordination, and construction administration. The work of GWWO Inc./Architects will be used as a basis for providing real world examples of how the program has been used through all stages of design and construction. The 70 000-sf² Daniel Z. Gibson Performing Arts Center at Washington University will be a featured project. The center provides a home for both the Music and Drama Departments and includes state of the art teaching and performance spaces. Challenges for the project included transforming an existing 600 seat house theater into a more intimate 440 seat theater with a balcony level, designing an acoustically pleasing 200 seat recital hall while minimizing the sound focusing of a concave glass wall along one side of the hall and the addition of a 175 seat flexible experimental theater.

8:15

5aAA3. Revit and the role of architectural acoustics in the design process. Jeff Larrick (Autodesk, West Palm Beach, FL 33401, jeff.larrick@autodesk.com)

A presentation by Autodesk discussing the role of Revit in the building information modeling design process and how it allows for collaboration between the architect and acoustical consultant.

Contributed Papers

8:35

5aAA4. Distilling the acoustical model from BIM (Building Information Modeling) Standard architectural, mechanical and structural models: Robust acoustical templates, limitations, and recommendations. Richard A. Vedvik and Jon W. Mooney (KJWW Eng. Consultants, 623 26th Ave., Rock Island, IL 61201, vedvikra@kjww.com)

The integration of architectural acoustics into building integration modeling (BIM) software and the implementation of such integration is discussed between available systems and platforms. The process of integrating the various platforms of building integration modeling software, in such a manner that the assigned properties can be shared, manipulated, and exported, is of great interest to acousticians. As the architectural model is created, a standardization of materials and generic construction methods is desirable so the relevant properties assigned to the various building components are able to be imported and adjusted by robust acoustical templates (RATs). Additional complications are identified as the acoustical model is detached from the live model. Standardization is also important between the architectural, structural, and mechanical models in order to create a common shared model that is able to maintain the relevant properties, both physical and analytical. Finally, the operation of a RAT prototype is demonstrated in the distillation of an acoustical model from standard BIM architectural, mechanical, and structural models.

8:50

5aAA5. Error estimation due to sample size effects of in situ surface impedance measurements. Eric Brandao, Erico Fulco, Arcanjo Lenzi (Federal Univ. of Santa Catarina, Joao Pio Duarte Silva St., 250, B13, Ap 201 Corregio Grande Florianopolis-SC, 88037-001 Brazil, eric@lva.ufsc.br), and Emiel Tijs (Microflown Technologies)

The boundary element method (BEM) is used to predict the effect of the sample size on in situ impedance measurements on a locally reactive sample
at normal incidence. The BEM model is based on a direct measurement of the acoustic impedance above the sample (with a combined sound pressure and particle velocity sensor). As the sound source is placed close to the sensor, the acoustic waves cannot be considered plane anymore. Because of that, two models to obtain the surface impedance based on the acoustic field are presented and compared. It is also shown that the sample size, the thickness of the sample, and the positions of the source and receiver affect the measured surface impedance and the calculated absorption coefficient.

Those effects are parameterized in a way that the user of the in situ measurement setup may have an estimate of the required sample size for a desired frequency range with a given error allowance.

9:05 5aAA6. Examining the relationship between total acoustic absorption and late lateral energy (GLL). David A. Dick and Michelle C. Vigant (Dept. Mech. Eng. Acoust. Prog. and Lab., Univ. of Hartford, 200 Bloomfield Ave., West Hartford, CT 06117, vigant@hartford.edu)

Listener envelopment, the sense of being immersed in the sound field, has been shown to correlate with the objective parameter late lateral energy (GLL). Barron proposed that GLL is primarily a function of the ratio of reverberation time to auditorium volume or the total acoustic absorption [Barron, Appl. Acoust. 62, 185–202 (2001)]. An investigation has been conducted to examine this proposed theory with measurement data and predicted values from a room acoustics computer modeling program. Room impulse response measurements have been taken in the Belding Theatre in downtown Hartford, CT. Features of the hall include variable absorption in the form of curtains, both within the hall and also within small adjacent coupled volumes. Measurements were taken in several absorption configurations. A detailed room acoustics computer model of the hall has been created in order v9.22 to examine the accuracy of the model’s prediction of GLL. The model was initially validated using the measured parameters of reverberation time (T30), early decay time (EDT), and clarity index (C80). A comparison of the measured and predicted results will be discussed, along with an analysis of the relationship between GLL and total acoustic absorption. [Work supported by a University of Hartford Greenberg Junior Faculty Grant.]

9:20 5aAA7. Classroom acoustics in the developing world: A student project to develop simple assessments and treatments. Nathan McNeill, Matthew Blevins, Emily Drott, Alex Kremer, Mark Kusch, Brian Pluha, Ashley Ringer, Valerie Sargent, and J. Stuart Bolton (Purdue Univ., Neil Armstrong Hall of Eng., Rm. 1300, 701 W. Stadium Ave., West Lafayette, IN 47907-2016, nmcmuell@purdue.edu)

An inadequate educational system can be crippling to a developing country struggling to advance in the modern world. Poor classroom acoustics are a significant factor that can contribute to an inadequate education. Classroom acoustics is a mature field of research in the developed world; however, the classroom conditions and resources found in the developing world present new research challenges. A special project course was developed at Purdue University to give undergraduate students the opportunity to learn about and address classroom acoustics problems in underdeveloped regions. The goal of this special project course was to develop an evaluation and design manual that can be distributed to teachers and educational officials in those regions. The purpose of the manual is to help identify acoustic problems in classrooms, provide suggestions of simple noise control solutions, and outline best practices for the design of new classrooms. The manual specifically addresses problems related to excessive reverberation, poor signal-to-noise ratios, and noise transmission from outside the classroom. The ultimate intent of this project was to increase awareness of the unique classroom acoustic conditions in the developing world.

9:35 5aAA8. Experimental validations of the transport equation model for room-acoustic predictions in long spaces. Yun Jing and Ning Xiang (Graduate Program in Architectural Acoust., School of Architecture, Rensselaer Polytechnic Inst., Troy, NY 12180)

This paper presents an experimental study on steady-state sound pressure levels and reverberation times in a tenth long-space scale model. The purpose of this work is to validate a recently developed one-dimensional transport equation model for acoustic predictions in long spaces. A series of measurements has been conducted in the scale model to reveal receiver-position dependence of measured and modeled parameters in the long spaces. Experimental results at different octave bands are obtained. This paper will compare the acoustically measured results with that of simulations.

9:50 5aAA9. An experimental scaled-model for coupled volumes with an automated high-spatial-resolution scanning system. Joonhee Lee, Craig Schaefer, and Ning Xiang (Graduate Program in Architectural Acoust., School of Architecture, Rensselaer Polytechnic Inst., Troy, NY 12180, gj20plus@gmail.com)

This paper presents an experimental study of low-frequency behavior in coupled volumes. In order to examine low-frequency behavior in coupled volumes, an eighth-scale model of two coupled volumes with an automated high-spatial-resolution scanning system has been developed. The validated scanning system makes it possible to acquire thousands of room impulse responses on a designated planar grid with a fine grid size in the scaled model. The high-spatial-resolution scanning results are used to investigate the characteristics of acoustic wave propagation in both single-space and coupled-volume systems. This paper discusses some design issues relevant to high-quality, high-spatial resolution scanning, to diffusely reflecting interior surfaces, and to variable coupling aperture sizes. Some representative scanning results of sound pressure propagations and sound energy flows will also be demonstrated.

10:05—10:20 Break

10:20 5aAA10. Some preliminary investigations on the use of a diffusion equation model for room-acoustic parameters prediction. José Escolano (Dept. Telecommunication Eng., Univ. of Jaén, 23700, Linares, Spain, escolano@ujaen.es), Juan M. Navarro (San Antonio’s Catholic Univ. of Murcia, 30107 Guadalupe, Spain), José J. López (Universidad Politécnica de Valencia, 46021 Valencia, Spain), and Ning Xiang (Rensselaer Polytechnic Inst., Troy, NY 12180)

Over the very recent years, an alternative model to describe sound fields in enclosures, based on the mathematical theory of diffusion, has been proposed. This model is based on the concept of sound particles traveling along straight lines and striking walls or scattering objects existing in a room. The scope of this work is to present some preliminary investigations in the prediction of objective room-acoustics parameters, such as reverberation time, clarity, and definition through the use of a diffusion equation. This paper will discuss modeling accuracy of the room-acoustics parameters using some numerical examples. Moreover, some issues concerning the numerical implementation by means of finite difference schemes for the diffusion equation are addressed.

10:35 5aAA11. Micro-perforated stretched foils as sound absorbers and barriers. Christian Nocke, Catja Hilge (Akustikbuero Oldenburg, Katharinenstr. 10, 26121 Oldenburg, Germany, nocke@akustikbuero-oldenburg.de), and Jean-Marc Scherrer (BARRISOL S.A.S, F-68680 Kembs, France)

A typical set-up of a micro-perforated sound absorber is a micro-perforated sheet or panel in front of a closed air cavity. The sound absorption coefficient of these set-ups can be easily calculated with a high accuracy according to the well-known approximation of D.-Y. Maa if all defining geometrical parameters (diameter of microperforation, distance between orifices, panel thickness, and air cavity depth) are known. Other applications are freely suspended sails of micro-perforated sheets as well as double layer structures mad of micro-perforated materials. In this contribution measured sound absorption coefficients of different set-ups as well as measurements of level decrease will be presented and discussed. For some of these assemblies
no closed calculation model exists so far. Comparisons between computer-based calculations of layered absorber set-ups show good results for some assemblies. Finally room acoustic applications of micro-perforated stretched sheets are shown.

10:50

5AA12. Insertion loss estimation of an opened enclosure using a statistical ray-tracing model. Buye Xu, Scott D. Sommerfeldt, and Timothy W. Leishman (Dept. of Phys. and Astronomy, Brigham Young Univ., Provo, UT 84606, buye.xu@gmail.com)

Estimating the insertion loss (IL) of an enclosure with one or more openings is important for many noise control applications. In the low-frequency range (below the Schroeder frequency) where modal effects are dominant, numerical simulations based on modal expansions or the finite element method are feasible for obtaining the desired predictions. Above the Schroeder frequency, statistical room acoustics theories are more desirable because numerical simulations are not applicable or are computationally expensive. However, existing room acoustics theories can only deal with simple cases where a diffuse field can be assumed. The acoustic field inside enclosures is generally sufficiently complicated in practice such that the diffuse field assumption is not valid. Examples of the complex conditions include narrow spatial regions, non-uniform boundaries, and the existence of obstacles in the enclosure. In this paper, a model that adopts the basic ideas of both ray-tracing and Sabine’s room acoustics theory will be introduced to deal with the difficulties mentioned. This new model greatly improves the ability to estimate acoustic performance when the enclosure conditions are not ideal, yet requires much less computation than the normal ray-tracing model. Comparisons of different models with experimental data will be presented.

11:05

5AA13. High-resolution absorption mapping with a pu surface impedance method. Emiel Tijs (Microflown Technologies, 6826 CC Arnhem, Arnhem, The Netherlands, tijs@microflown.com)

The in situ method to measure the surface impedance with a pu probe documented is published in many publications. The method is based on the measurement of sound pressure (p) and particle velocity (u) close to an acoustic absorbing material. A loudspeaker at a defined distance is applied to generate a sound field with a known radiation impedance. The impedance of a small area (a few square cm²) with a known impedance is scanned with an ultraminiature pu probe very close to the surface. The area is made of steel with a cut-out, behind this a material with a known impedance is placed. In this paper the method is explained, the spatial accuracy of the measurement is examined, and a visualization technique is presented with a display of the spatial distribution (two dimensional picture) of the damping properties as function of frequency. Finally the high-resolution measurements are combined and compared with a method that covers a larger surface area.

11:20

5AA14. Is it 6 decibels? The effect of pressure doubling on exterior to interior transmission loss values. Noel Hart, Kimberly Lefkowitz, Arno S. Bommer, and Robert D. Bruce (CSTI Acoust., P.O. Box 218808, Houston, TX 77218-8808, noel@cstiacoustics.com)

When calculating sound levels inside a building due to external sources, should some value be added to the expected sound level? ATSM E966 recommends adding 6 dB to this level. It has been theorized that this is due to pressure doubling. The effect of pressure doubling on the calculated exterior to interior sound transmission in the direct field is, however, often misunderstood. Most texts on the subject simply list a 5- or 6-dB correction factor that is applied to the transmission loss value. The reasoning of this particular number is neither expounded upon nor explained. This paper investigates and explains this correction factor.

11:35—12:00 Panel Discussion

FRIDAY MORNING, 23 APRIL 2010

GRAND BALLROOM I/II, 8:45 A.M. TO 12:00 NOON

Session 5aAB


Ann E. Bowles, Cochair

Hubbs Sea World Research Inst., 2595 Ingraham St., San Diego, CA 92109

Sean K. Lehman, Cochair

Lawrence Livermore National Lab., Livermore, CA 94550

Chair’s Introduction—8:45

Invited Papers

8:50

5aAB1. Multipath features as a contextual clue for classification. Christopher O. Tiemann (Appl. Res. Labs., Univ. of Texas at Austin, P.O. Box 8029, Austin, TX 78731)

A single odontocete click is often heard as multiple click echoes due to the many paths sound travels between the source (whale) and receiver; these echoes are called multipath arrivals, and together they form an arrival pattern. Because each arrival can appear differently due to propagation effects, efforts to associate recorded clicks with individual whales can be frustrated by the presence of multipath arrivals, particularly when many whales click simultaneously. Automated classifiers may associate arrivals from one click event into many groups, even though they come from one source, thus overestimating the number of whales in a recording. However, collectively a multipath arrival pattern can be exploited to huge advantage when trying to do individual association. Typically, the relative spacing
of arrivals changes slowly over time (due to slow movement of the source) and is unique for a given source location. The persistence of stable arrival patterns over successive click events provides a contextual clue for associating many arrivals to a single individual or can validate automated classification techniques. Finding unique arrival patterns should provide an accurate estimation of the number of sources in an area and even serves as the first step in passive acoustic tracking of whales.

9:10

5aAB2. Contextual rhythmic analysis of beaked whale clicks for passive acoustic identification. Natalia Sidorovskaia, Philip Schexnayder (Dept. of Phys., Univ. of Louisiana at Lafayette, UL Box 44210, Lafayette, LA 70504-4210, nas@louisiana.edu), George E. Ioup, Juliette W. Ioup (Univ. of New Orleans, New Orleans, LA 70148), Christopher O. Tiemann (Univ. of Texas at Austin, Austin, TX 78713), Alexander E. Ekomov, and James Sabatier (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, Oxford, MS 38677)

The Littoral Acoustic Demonstration Center (LADC) has collected an extensive acoustic library of marine mammal phonations in the Northern Gulf of Mexico, including the Cuvier’s and Blainville’s beaked whale clicks. An automatic algorithm, which performs cadence frequency analysis, has been developed and successfully applied to sperm whale clicks to identify temporal evolution of individual rhythmic patterns in overlapping trains of simultaneously clicking whales in a group. In this study the results of the association of cadence frequencies with individual beaked whales in a group for the LADC 2007 Gulf of Mexico recordings are reported. The algorithm is a part of an integrated tool proposed by LADC for identification of individual animals in a group. The cadence frequency association results are compared to multi-attribute self-organizing map clustering and to time-of-arrival analysis and localization. A multi-band version of the algorithm is also presented. It provides species association cues for complex cluttered acoustic environments where different species phonate with similar inter-click intervals. The proposed approach is beneficial for passive acoustic studies of beaked whale populations and social behavior and may contribute to understanding acoustic communication of these evasive animals. [Research supported by ONR and SPAWAR.]

9:30

5aAB3. Using click change detection to modify cluster analysis for acoustically identifying individual sperm whales with changing aspect. George E. Ioup, Juliette W. Ioup, Lisa A. Pflug (Dept. of Phys., Univ. of New Orleans, New Orleans, LA 70148, geioup@uno.edu), Christopher O. Tiemann (Univ. of Texas at Austin, Austin, TX), and Natalia A. Sidorovskaia (Univ. of Louisiana at Lafayette, Lafayette, LA)

As described by Mohl et al. J. Acoust. Soc. Am. 114, 1143–1154 (2003), the acoustic properties of the received clicks of a turning sperm whale change as the whale rotates. These changing properties (the time signal, the discrete Fourier transform, and the discrete wavelet transform are studied here) confound cluster analysis attempts to identify individual whales by grouping like clicks in classes representing individuals. The 4th International Workshop on Detection, Classification and Localization of Marine Mammals Using Passive Acoustics has associated workshop data containing a sequence of sperm whale clicks which most likely come from the same individual (as shown by careful manual click analysis) turning with respect to the acoustic recording buoy. The time domain sequence of clicks indicates that the angle of a line from the head of the whale to the buoy appears to be changing from greater than 90 deg to less than 90 deg and back again. Time evolution and click change detection are applied to follow the whale and give correct automatic identification. This analysis is incorporated into the clustering process. [Research supported by ONR and SPAWAR.]

9:50

5aAB4. Acoustic identification of individual beaked whales using self-organizing maps for automated classification of clicks in Littoral Acoustic Demonstration Center (LADC) data. Juliette W. Ioup, George E. Ioup, Lisa A. Pflug (Dept. of Phys., Univ. of New Orleans, New Orleans, LA 70148, jioup@uno.edu), Natalia A. Sidorovskaia, Philip Schexnayder (Univ. of Louisiana at Lafayette, Lafayette, LA), Christopher O. Tiemann, and Alan Bernstein (Univ. of Texas at Austin, Austin, TX)

The Littoral Acoustic Demonstration Center (LADC) acquired underwater acoustic data from clicking sperm and beaked whales in the Gulf of Mexico during an experiment in July 2007. Analysis of sperm whale echolocation and coda clicks in the data has been presented at previous ASA meetings. Research emphasis is now on identifying individual beaked whales from their echolocation clicks. Again, clustering methods present one appropriate approach to the automated identification of these whales by computer. Self-organizing maps have the ability to limit the number of clusters automatically, and their successful use with sperm whales encourages the application to beaked whales as well. Results for time, frequency, and wavelet domain data allow comparison among clustering done using different signal representations and help highlight areas of concern in the analysis. As in the case of sperm whales, the beaked whale identifications are verified by comparing to cadence analysis (as presented in an accompanying abstract) and to a limited extent by manual analysis and localization. [Research supported by ONR and SPAWAR.]

10:10—10:25 Break
Contributed Paper

10:25

5aAB5. Automated detection of sperm whales (Physeter macrocephalus) creaks and association with depredation events. Delphine Mathias, Aaron Thode (Marine Physical Lab., Scripps Inst. of Oceanogr., 9500 Gilman Dr. La Jolla, CA, 92037-0238, delphine.mathias@gmail.com), Jan Straley (Univ. of Alaska Southeast, Sitka, AK 99835), John Calambokidis, Gregory S. Schorr (Cascadia Res. Collective, Olympia, WA, 98501), and Bill Burgess (Greeneridge Sci. Inc., Goleta, CA 93117)

Sperm whales have been depredating black cod (Anoplopoma fimbria) from demersal longlines in the Gulf of Alaska for decades, but the behavior has now become pervasive enough that it is starting to affect government estimates of the sustainable catch, motivating further studies of this behavior. In July 2007 and June 2009, bioacoustic tags were attached to adult sperm whales off Southeast Alaska under both natural foraging conditions and situations wherein the animals were depredating sablefish from commercial longlining vessels. Tags remained attached for a cumulative 100 h of depth, orientation, and acoustic data. Results show that when depredating, sperm whales often dive at depths shallower than 50 m, compared to natural foraging depths of 300–400 m in the area. Dive durations, pitch, and roll rates are also distinctively different. Techniques such as spectrogram cross-correlation and cepstral analysis are used to automatically detect creak events, which have lower intensities and narrower frequency ranges than conventional clicks. Acoustic measurements such as creak production rate, duration, and intensity are then extracted. Correlations are derived between creak events, sudden changes in body orientation and depredation rate, with a goal of determining an acoustic measure of depredation activity and efficacy. [Work conducted under the SEA SWAP Program, supported by the North Pacific Research Board and the National Geographic Society.]

Invited Paper

10:40

5aAB6. Detecting sequences of calls. David K. Mellinger (Coop. Inst. for Marine Resources Studies, Oregon State Univ., 2030 SE Marine Sci. Dr., Newport, OR 97365, david.mellinger@oregonstate.edu)

Many sound-producing animals make vocalizations with predictable timing patterns. In some cases, animals produce long, regular sequences of sounds, either for breeding (e.g., many baleen whales) or for foraging (e.g., most toothed whales and dolphins); in other cases, different call types are produced in succession with the intervals between calls following certain timing patterns. Here I describe methods for detecting these timing patterns. Both of these methods operate on a “detection function,” the estimated likelihood at each instant of time that a call is present. Long, regular sequences can be detected using an autocorrelation of the detection function that significantly increases the detectability of sequences compared to single calls. Different call types occurring in irregular but predictable timed patterns can be detected using methods similar to dynamic time warping; again, the detectability of the patterns is notably higher than the detectability of single calls. [Work funded by ONR and the Navy’s Environmental Readiness Division.]

11:00—12:00 Panel Discussion
5aBB2. Quantitative ultrasound in the management of osteoporosis, Marc-Antoine Krieg and Hans Didier (Ctr. of Bone Diseases, CHUV, Av. Pierre-Decker 4, 1011 Lausanne, Switzerland)

Bone mineral density (BMD) is one of the major determinants of bone strength and fracture risk (FR), but it is not the only determinant, and overlap exists in BMD values between patients with and without fractures. It is therefore important to improve the strategy in better detecting patients at high FR who will benefit from treatments (TT). Quantitative bone US(QUS) are portable and less expensive than DXA. Only few QUS measuring the heel have reached a sufficient level of validation to be routinely used. They usually measure the attenuation (BUA) and the velocity (SOS) of the US wave. It is well established that valid heel QUS predict fragility fracture in PM women and men >65, independent of BMD. Then, heel QUS can be used in combination with other clinical risk factors (CRFs) in a model of risk assessment. Then, when DXA is available, the primary clinical use of heel QUS is to identify, in conjunction with CRF, a population at low FR in which no further diagnostic evaluation is necessary. On the other hand, if central DXA cannot be done, pharmacologic TT can be initiated if the probability, as assessed by heel QUS in conjunction with CRF, is sufficiently high.


Developments in hardware and software capabilities, along with a significant increase in access to equipment and awareness of the methods, over the last 10 years have enabled micro-computed tomography (CT) technology to establish itself as one of the basic tools to study bone micro-architecture in the laboratory. A wide range of equipment is now available ranging from synchrotron based facilities to compact desktop scanners for small to large specimens and scanners for in vivo pre-clinical imaging. The various scanners offer the ability to acquire images with resolution ranging from sub-micron to 100 μm in acceptable scan times and where applicable, with acceptable x-ray doses. More recently, high-resolution scanners (under 100 μm) have become available for clinical research. Micro-CT is now widely used to investigate bone from pre-clinical models of various aging and pathologic conditions (osteoporosis, osteoarthritis, and arthritis) as well as models for tissue regeneration and bone fracture repair as well as high-throughput skeletal phenotyping. Pre-clinical in vivo imaging methods are now routinely used as a screening tool in drug discovery and development processes. The availability of sophisticated software for analyses and visualization of these images has made micro-CT a very versatile research tool for the pre-clinical laboratory and promises to do the same for clinical research.

9:00


5aBB4. Bone diagnostics using dual frequency ultrasound measurements, Jukka S. Jurvelin, Janne Karjalainen, Ossi Riekkinen, Markus Maño, Hanna Isaksson, and Juhu Týräs (Dept. of Phys. and Mathematics, Univ. of Eastern Finland, P.O. Box 1627, 70211 Kuopio, Finland, jukka.jurvelin@uef.fi)

Diagnosis of osteoporosis is made at skeletal sites composed mainly of trabecular bone. Cancellous has been the first location for through transmission (TT) ultrasonic measurements of trabecular bone. Similarly as with the DXA, the best prediction of the hip fractures could be obtained by making the measurements at hip. To realize axial pulse-echo (PE) ultrasound measurements, e.g., in hip we have introduced the dual frequency ultrasound (DFUS) technique to minimize the effects of soft tissues overlying the bone and to correct the PE-parameters [Riekkinen et al., Ultrasound Med. Biol. 34, 1703–1708 (2008)]. Based on our experimental measurements at frequencies of 2.25 and 5.0 MHz and finite difference simulations, the DFUS technique detects minor changes in bone density, despite variable composition of soft tissue. For TT-geometry, the composition of the interfering soft tissues may also be solved with the DFUS, e.g., by measuring the reflection from the bone surfaces at both sides of the bone. The measurements with phantom materials indicated that accuracy of TT-parameters, such as broadband ultrasound attenuation, were significantly improved by DFUS. Potentially, the DFUS could also benefit the in vivo TT-measurements of trabecular bone. [Support from Academy of Finland is acknowledged.]

Contributed Papers

9:30

5aBB5. Comparison of conventional ultrasonic phase velocity and attenuation measurements of cancellous bone to estimates obtained using Bayesian probability theory, Christian C. Anderson (Washington Univ. in St. Louis, 1 Brookings Dr., St. Louis, MO 63130), Michal Pakula (Kazimierz Wielki Univ., Bydgoszcz, Poland), Pascal Laugier (Université Pierre et Marie Curie, Paris, France), G. Larry Brethorst, Mark R. Holland, and James G. Miller (Washington Univ. in St. Louis, St. Louis, MO)

Cancellous bone supports the propagation of two compressional waves, commonly referred to as fast and slow waves, which can overlap in the acquired time-domain data. When such data are processed using conventional techniques, interference effects may cause artifacts to be present in the speed of sound and broadband ultrasound attenuation (BUA) measurements. Bayesian probability theory is a proposed method for extracting the ultrasonic properties of the individual fast and slow waves, even in the presence of substantial interference. In the current study, data were acquired in vitro at sites on a cancellous bone, and the acquired signals were used as input to a Bayesian probability approach for estimating parameters characterizing the fast and slow waves. For comparison, the signals were also processed using conventional methods. Although the conventionally obtained phase velocity and BUA frequently exhibited anomalous features, including negative dispersion, those estimated using probability theory exhibited no aberrant behavior. Hence, material properties obtained using Bayesian probability theory may provide more reliable estimates of bone quality and fracture risk than the apparent properties derived from the non-causal mixed mode signals. [Work supported by NIH R01HL40302 and R01AR057433, NSF CBET-0717830.]

9:45

5aBB6. Transfer functions for one-dimensional bone models, James L. Buchanan (Dept. Mathematics, US Naval Acad., 572C Holloway Rd., Annapolis, MD 21401, jlb@usna.edu), Robert P. Gilbert (Univ. of Delaware, Newark, DE 19711), and Miao-Jung Ou (Oak Ridge Natl. Lab., Oak Ridge, TN 37831)

A system for calculating wave pulses through one-dimensional layered media via a series of transfer functions is developed. This system obviates the need for numerical approximation beyond the use of discrete Fourier transforms, whence it is not necessary to deal with the poor numerical conditioning of the linear systems that arise in multi-layered models. A one-dimensional model of bone as a cancellous segment between two cortical segments isonified by an ultrasonic pulse is considered. For the cancellous
The governing equations for wave motion in the linear theory of anisotropic poroelastic materials have been developed and extended to include the dependence of the constitutive relations upon fabric. Fabric is a quantitative stereological measure of the degree of structural anisotropy in the pore architecture of a porous medium. This anisotropic poroelastic theory is then applied to cancellous bone architecture to study the role of both porosity and fabric on the wave attenuation due to reflection, absorption, and scattering. The analysis of the fast and slow wave attenuation is performed as a function of the porosity and fabric of the trabecular structure. Wave attenuation predictions from this model confirm our previous experimental observations indicating that either the fast or slow wave may be the predominant wave mode in trabecular bone depending on the bone architecture.

The analysis of tortuosity is a function of the fabric tensor. This relationship between tortuosity and fabric is consistent with the presentations of Biot and later authors. In this work we show that tortuosity is a function of the fabric tensor. This relationship between tortuosity and fabric is employed to provide a geometric understanding of tortuosity in terms of pore space architecture. This geometric understanding is then applied to cancellous bone architecture. This work was supported by grants from NSF (NSF/CUNY AGEP 0450360, NSF-MRI 0723027, and PHY-0848491). The authors also acknowledge the support from PSC-CUNY Research Award Program (PSC-CUNY-60014-38-39, 69486-00 38, and PSCREG-39-1103) and NIH 1SC2AG34198-1A1.

This investigation addresses the inverse problem of recovery of poroelastic parameters of cancellous bone by inversion of transmitted ultrasonic data. Zine El Abidine Fellah, Erick Ogam (Laboratoire de Mecanique et d’Acoustique, UPR 7051 CNRS, 31 Chemin Joseph Aiguier, 13402 Marseille, France), Catherine Masson (INRETS-UMR T 24 LBA UFR de Medecine, 13916 Marseille Cedex 20, France), and Robert Gilbert (Univ. of Delaware, Newark, DE 19716)

This investigation addresses the inverse problem (IP) of the retrieval of the macroscopic poroelastic parameters of cancellous bone using transmitted ultrasonic waves. The IP is solved by minimizing the objective functional (OF), which is a measure of the discrepancy, in the least squares sense, between the trial and measured data pertaining to the transmitted USW. In this study the Biot–Johnson–Koplik was employed as the interaction model (IM). A close examination of the OF computed using this IM and measured ultrasonic data transmitted through the cancellous bone showed that it exhibited several local minima and maxima. Most off-the-shelf iterative non-linear least squares optimization algorithms (such as the Levenberg–Marquardt or Nelder–Mead simplex methods) could not retrieve a nonambiguous solution. A simple method for resolving this problem without ambiguity by using two distinct transducer central frequencies is proposed. Finally the recovered acoustic parameters are compared with those obtained using classical methods.

If dispersion in a medium is weak and approximately linear with frequency (over an experimental band of frequencies), then it can be shown that the constant term in a polynomial representation of phase shift as a function of frequency can produce errors in measurements of phase velocity differences in through-transmission, substitution experiments. A method for suppressing the effects of the constant phase shift in the context of the single-wave-model was tested on measurements from 30 cancellous human calcaneus samples in vitro. Without adjustment for constant phase shifts, the estimated phase velocity at 500 kHz was 1516 + 6 m/s (mean + standard error), and the estimated dispersion was –24 + 4 m/Hz (mean + standard error). With adjustment for constant phase shifts, the estimated mean velocity decreased by 4–9 m/s, and the estimated magnitude of mean dispersion decreased by 50%–100%. The average correlation coefficient between the measured attenuation coefficient and frequency was 0.997 + 0.0026 (mean + standard deviation), suggesting that the signal for each sample was dominated by one wave.

In this paper, we determine bone fragility of cancellous bone in vivo by using a simple one-dimensional model of a muscle-cortical bone-cancellous bone sample, by measuring the refracted sound when the sample is exposed to impulses from a transducer. From these data, the bone parameters can be determined. Such a procedure is referred to as solving an inverse problem. A transducer is placed on one side of the skin of the member to be interrogated and a receiver is place diametrically opposed in contact with the member. The muscle will be modeled as an elastic material, whereas the cortical and cancellous bone may be modeled by modified Biot equations albeit with different porosities. However, because of the known frequency dependence of the viscous term, we propose that the standard Biot equations, which are not well posed, be replaced by adding a convolution term. The form of the kernel \( k(t) \) in the convolution integral is suggested by frequency domain representation of the dissipation term in the Biot–Johnson–Koplik–Dashen model.

The semi-empirical scattering models of trabecular bone were developed and examined for their abilities to mimic the frequency dependent backscattering coefficient measured in the cancellous bone. In the simulation of the bone, rf echoes the real properties of the bone and experimental conditions were taken into account. Tree types of trabeculae mimicking scatterers were considered. First, the bone consisted of cylinders with varying thickness (Gammadistributed) within the population was assumed. The next two cases accounted for the contribution of thick and thin trabeculae to the total backscattered signal. The second model assumed existence of two populations of the cylindrical scatterers significantly differing in the average value of Gamma distributed diameters. Finally, the mixed model composed of thick and thin trabeculae modeled, respectively, by cylindrical and spherical scatterers was examined. The last selection resulted from the similarity found between scattering on small sphere and finite cylinder. Calculated echoes demonstrated the usefulness of the mixed model. Frequency dependence of backscattering coefficient agreed well with the experimentally de-
Quantitative ultrasound imaging is a potential method for the evaluation of the integrity of articular cartilage. A quantitative minimally invasive intrarticular ultrasound (IAUS) technique has been developed for in situ and in vivo diagnostics of articular cartilage [Virén et al. Ultrasound Med. Biol. 35, 1546–1554, (2009)]. In IAUS, a high-frequency (40 MHz) intravascular ultrasound device is used and reflection coefficient ($R$), integrated reflection coefficient (IRC), apparent integrated backscatter coefficient (AIB), and ultrasound roughness index (URI) are calculated for cartilage. A set of experiments was conducted for the validation of the technique. For example, significant differences were observed in ultrasound parameters between intact, spontaneously and surgically repaired rabbit articular cartilage. The parameters revealed abnormal surface roughness and internal collagen structure by decreased $R$ and IRC and increased values of URI and by increased AIB. Furthermore, the poor integration of the repair and surrounding native tissue was visible in the ultrasound images. As a reference, histological images revealed that repaired tissue exhibited abnormally organized collagen network, rough surface and lower collagen, and proteoglycan contents. Based on the present experimentation using IAUS, tissue structure and properties, not diagnosable by present clinical techniques, may be quantitatively evaluated with the minimally invasive ultrasound technique.

FRIDAY MORNING, 23 APRIL 2010

Session 5aEA

Engineering Acoustics: Electret Condenser Microphones

Allan J. Zuckerwar, Cochair
4909 Camberley Cir., Williamsburg, VA 23188

Qamar A. Shams, Cochair
NASA Langley Research Ctr., 4 Langley Blvd., Mail Stop 238, Hampton, VA 23681

Chair’s Introduction—8:00

Invited Papers

8:05

5aEA1. Electret condenser microphones. James E. West (Dept. of Elec. and Comput. Eng. and Dept. of Mech. Eng., Johns Hopkins Univ., 3400 N. Charles St., Baltimore, MD 21231, jimwest@jhu.edu)

For more than 4 decades electret microphones have dominated most transducer applications in the audio frequency range for three main reasons, linearity, simplicity, and they are inexpensive. E. C. Wente invented the condenser microphone in 1916 searching for a replacement for the carbon microphone used in most telephones, but it was impractical until the high-voltage battery was replaced with a permanently charged polymer. Electret microphones have been the workhorse for most microphone applications including microphones for professional sound level measurements, telephones, hearing aids, professional audio recordings, games, and toys. Recent worldwide estimates of the electret microphone market suggest that over 2 billion units are made annually. Electret condenser microphones are available in a variety of sizes from very small (a few millimeters in diameter) to about 2 in. in diameter. The use of multiple microphone elements in arrays and the use of digital signal processing provide interesting prospects for improving signal-to-noise. A variety of microphone arrays and some important historical facts will be discussed in this talk.

8:25

5aEA2. Piezoelectret microphones. Gerhard M. Sessler and Joachim Hillenbrand (Dept. of Elec. Eng., Univ. of Technol., Merckstrasse 25, 64283 Darmstadt, Germany, g.sessler@t.tu-darmstadt.de)

Piezoelectret microphones, a recently described new class of electret microphones, are based on the strong piezoelectric effect of internally charged, cellular polymer films. Such piezoelectret films have a large number of internal voids within the polymer. The upper and lower walls of these voids are oppositely charged. Films of this kind, when metallized on both sides, show high-piezoelectric activity and can be directly used as microphones. Recent work on these transducers has concentrated on stacking of piezoelectret films to improve sensitivity, on directional microphones, and on investigations of new film materials with improved charge stability at elevated temperatures. Stacked transducers have sensitivities proportional to the number of piezoelectret layers. With six layers, values of approximately 15 mV/Pa can be achieved. Directivities include omnidirectional, bidirectional, and cardioid patterns as well as dimensional characteristics. The new materials are based on fluoropolymers. These are, as opposed to the originally used polypropylene, not available in suitable cellular form. However, useful films can be made by fusing thin fluorocarbon films such that an array of small air voids is created between the layers. Transducers consisting of charged films of this kind withstand temperatures of 90 °C without significant loss of sensitivity.
5aEA3. Temperature characteristics of electret silicon microphones. Yoshinobu Yasuno (Appl. Piezoelectricity Div., Kobayasi Institute of Physical Res., 3-20-41, Higashi-motomachi, Kokubunji, Tokyo 185-0022, Japan, yasuno@kobayasi-riken.or.jp)

The electret condenser microphone (ECM) has become an important component for various consumer equipment systems because of its stable sensitivity and frequency characteristics and its success in achieving small size and high sensitivity. A previous report described the method to design a microphone with stable temperature characteristics. The quality of the silicon material for the diaphragm was pointed out as important for improving robustness. Recently, microphones using micro-electro-mechanical systems (MEMS) have been practically applied and widely used for mobile equipment, such as cellular phones. The major reason for adopting MEMS is that a re-flow soldering process can be used in production, in addition to having desirable features such as being small and thin. However, almost no electret types have been commercialized since a guaranteed 300°C heat resistance is required. This report presents a new trial electret silicone microphone that contains a SiO₂ thin-film electret. The report also evaluates the heat resistance of a silicon diaphragm, and the temperature change in sensitivity and frequency characteristics compared with those of a conventional ECM and the microphone for measurement.

9:05


30 years ago, Bruel & Kjaer launched the world’s first prepolarized condenser measurement microphones. These microphones combined the high-performance construction of the normal measurement microphones with a polarized electret layer placed on the back electrode. Since then, prepolarized microphones have gained an increasing share of the measurement microphone market. Prepolarized microphones are flexible in use because they do not need the external polarization, in particular, they can be used with constant current supply preamplifiers (Deltatron, ICP) that are widely used for many purposes in the industry. In order to be used for measurement purposes, the microphones and thereby the electret polarization must be highly stable and reliable. That is indeed achievable, and that is the main emphasis of this paper. A short historical overview of the development of the Bruel & Kjaer prepolarized microphone types is given. The properties of prepolarized microphones are discussed and compared to those of externally polarized microphones. These will include stability with respect to time, temperature, humidity and the chemical environment, as well as distortion and dynamic range.

9:25

5aEA5. Electret condenser microphones in hearing aids. Daniel Warren, David Schafer, and Kiran Konde (Knowles Electron., 1151 Maplewood Dr., Itasca, IL 60143, daniel.warren@knowles.com)

As sensors, small, high-gain devices, hearing aid microphones must be tiny, have high signal-to-noise ratios, be relatively insensitive to vibration from the nearby speaker, and operate at low hearing aid battery voltages. The subminiature backplate-electret condenser microphone was developed in the early 1970s specifically for hearing aids, with a very lightweight mechanical system for lower vibration sensitivity and somewhat improved signal-to-noise compared to previous designs. Although the design has been refined over several generations, the basic technology of electret films has proven to be very robust over the last 40 years, enabling continual size reductions while maintaining high signal-to-noise ratios. Noise generation and vibration sensitivity continue to be competitive dimensions for hearing aid microphones. Some of the mechanisms for noise generation and vibration pickup are intrinsic to the design of these microphones. Other mechanisms may be caused by interaction between the microphones and the hearing aid system of which they are a part. In this presentation, we will discuss both types of mechanisms and some design guidelines for minimal noise application of microphones in hearing aids.

9:45—10:00 Break

10:00

5aEA6. Ringing the changes: The potential for push-pull electret transducers in cell phones. Tim Mellow (Devices R & D, Nokia, Summit Ave., Farnborough GU14 0NG, UK, tim.mellow@nokia.com) and Leo Kärkkäinen (Nokia Res. Ctr., C00180 Helsinki, Finland)

Electret microphones are very common in mobile devices, but are invariably single-ended in configuration. For moderate dynamic conditions this is not a problem, but under high sound pressure levels, better linearity can be obtained using a push-pull arrangement. Recent progress in nano-porous electret materials also opens up possibilities for electret loudspeakers. Such loudspeakers would offer potential benefits such as high-conversion efficiency, no magnetic field, and an ultra-thin package. However, there are product integration challenges. The authors describe a simulation model of a push-pull electret transducer and discuss the conditions under which it is linear. It turns out that the relationship with a non-electret electrostatic transducer is relatively simple.

10:20


An extensive collection of airframe noise flight test data has been acquired for a regional jet class of aircraft using a ground-based electret condenser microphone phased array system. The measurements were conducted at the Wallops Flight Facility for the NASA Aeronautics Research Mission Directorate Fundamental Aeronautics Program. Gulfstream G450 and G550 aircraft were used as the test beds for the study. The phased array employed for the measurements consisted of a spiral clustering of 167 Panasonic WM-61A quarter-inch, omni-directional electret microphones. The innermost 49 sensors of the array were flush mounted on an aluminum plate with the remaining sensors flush mounted on individual 16-in. circular Plexiglas platters. The aluminum plate and platters were deployed on the overrun area of runway 4 at Wallops. Airframe noise measurements were performed with the aircraft flying a racetrack pattern simu-
lating the approach path to landing as it passed over the array. A distributed data acquisition system was used to acquire simultaneous
time histories from the array and was synchronized with instrumentation onboard the aircraft. This presentation will discuss the design
of the phased array, the calibration and deployment of the microphones, a sampling of the measurements obtained from the system, and
lessons learned from the deployment.

10:40

5aEA8. A method for free-field transfer calibration of electret microphones at ultrasound frequencies in air using a stable broad
band ultrasound source. Angelo J. Campanella (Campanella Assoc. and Acculab, 3201 Ridgewood Dr., Hilliard, OH 43026,
a.campanella@att.net)

A rotating turbulence screen wheel sound source, now in use in the air ultrasound instrument industry for product quality control,
provides a steady source of broad band ultrasound. This source is broad band through at least 100 kHz. It is calibrated as a secondary
standard, using a one-quarter inch condenser microphone, grid removed, normal incidence, and corrected for free-field conditions per
the microphone manufacturer free-field response. Sound pressure amplitude is measured at a distance of 500 mm and from 1 to 100 kHz
in 1-kHz FFT bins. Calibration site sound absorption by humid air is calculated to correct the emitted sound pressure values to be
reinserted for the user test site air temperature and humidity. Transfer calibrations of two electret condenser microphones are shown.
Microphone phase response measurement methods will also be investigated and discussed.

Contributed Paper

11:00

5aEA9. Acoustic field of air-coupled transducers made with
electroactive polymer foams. Louis Satyanarayan and Yves Berthelot
(Woodruff School of Mech. Eng., Georgia Tech Lorraine, GT-CNRS UMI
2958, 2 rue Marconi, 57070 Metz, France)

Electroactive films made of soft cellular polypropylene (PP) foams con-
taining electric charges have high piezoelectric coupling coefficient and
good matching impedance with air over a broad frequency range. They offer
interesting capabilities in the design of new air-coupled ultrasonic
transducers. A prototype transducer has been designed and tested in our
laboratory. The amplitude displacement on the surface of the transducer was
measured by laser vibrometry by vibrating the PP foams with a continuous
wave sinusoidal excitation. The transducer was mounted with a needle-type
PVDF hydrophone on a high-precision polar C-scan apparatus. The receiv-
ing hydrophone was translated along and across the axis of the source trans-
ducer and data were recorded. A quantitative characterization of the acoustic
field of the transducer was carried out by plotting the on-axis and cross-axis
variations in the pressure field, at various separation distances between the
source and the receiver in pulsed transmit-receive mode. The results indicate
that simple and relatively efficient air-coupled ultrasonic transducers can be
made with electroactive polymer foams. [Work supported by CNRS and by
the Conseil Rgional de Lorraine.]
5aMU2. Pitch bending and higher-mode reed vibration in mechanically blown free reed instruments. James P. Cottingham (Dept. Phys., Coe College, 1220 First Ave. NE, Cedar Rapids, IA 52402, jcotting@coe.edu)

Pitch bending in the harmonica, in which manipulation of vocal tract resonances plays an essential role, has long been a common practice. Pitch bending in free reed instruments with mechanically driven air supplies, such as the reed organ and the accordion, is a different matter. At least two methods of pitch bending in such instruments have been studied and demonstrated. The first, likely to be musically useful only in certain special circumstances, involves partial opening of the pallet valve combined with variations in blowing pressure. The second, recently described and implemented in the accordion by Tonon [J. Acoust. Soc. Am. 126, 2217 (2009)], involves modifying the construction to include a resonating chamber in addition to the standard reed chamber. It is also possible to construct air-blown free reed instruments in which the reed vibrates mainly in the second transverse mode. In one case the reed is coupled with a pipe resonator that provides a suitable mismatch in frequencies with the fundamental frequency of the reed. In another case the second or higher-mode vibration of the air-driven reed can be mechanically induced. The musical possibilities of these hypothetical reed-pipe instruments have yet to be explored.

9:00

5aMU3. Extended techniques of the contrabass. Michael T. Bullock (15 Channel Ctr. St., Apt. 403, Boston, MA 02210, michaelbullock@gmail.com)

Contrabass improvisers, composers, and performers of contemporary contrabass repertoire rely on a broad vocabulary of extended contrabass techniques. This vocabulary is continuously developed primarily through a kinesthetic engagement with the acoustics of the contrabass, and many new discoveries are made in the act of playing. This presentation will include live performance, by the presenter, of a series of extended techniques on the contrabass. Spectrograms and video clips will be used to illustrate how these techniques are related to the unique acoustic properties of that instrument. Principles demonstrated will include the following: (1) Overtone bowing and fingering. The low fundamental register of the contrabass means that each string has a many upper overtones within the audible range. Using various bowing and fingering techniques that exploit these upper partials, the timbre and pitch of a string can be radically altered and continuously transformed while playing. (2) Activating the contrabass body. Striking the contrabass body with hands and soft objects, rubbing the instrument’s back with the fingers, and manipulating feedback through amplification are ways to activate the complex resonances of the contrabass. [Thanks to Professor Jonas Braasch of the RPI School of Architecture.]

9:20

5aMU4. Extending and abstracting sitar acoustics in performance. Curtis R. Bahn (Dept. of the Arts, Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180-3590, ctb@rpi.edu)

There is a very close relationship between the acoustic design of instruments of the Hindustani classical music tradition and the structure of raga. The sitar and other related string instruments have sympathetic strings which are tuned to prominent notes of a raga. These strings resonate sympathetically in performance, creating a rich and hauntingly beautiful sonority. In creating a computer extended electronic sitar, the first step was to model and extend this resonant quality of the instrument, allowing for new tunings and extended resonances not possible with a traditional instrument. The extended sitar also uses performance analysis and a variety of sensors to control various signal processing techniques that further extend the quality of the instrument. This sonic analysis and subsequent sound synthesis/computer generated compositional schemes relate directly to structures of raga. This presentation will detail the development and performance of the computer extended sitar in relationship to traditional raga structure and in new contexts not directly related to traditional notions of the sitar.

9:40

5aMU5. Simulating, mirroring, morphing, and blurring acoustic space as an extended instrumental technique. Pauline Oliveros (Dept. Arts, Rensselaer Polytechnic Inst., 110 8th St., Troy NY 12080)

The paradigm for musical performance is the sound of musical instruments in a fixed acoustical space. Ordinarily space does not shift and change during performances. An ideal space for performing classical composers such as Mozart is the Shaffy Theater in Amsterdam with a relatively short reverberation time 1.5 s. Instruments couple to the space, and space is part of the instrument and vice versa. Players develop special relationships with favorite spaces. After experiencing many different types of acoustics this leads to the desire to change the space during performance in order to hear the different qualities that emerge from spatial coupling. Spatial progressions can be made so that rather than moving sound in space, space can be moved around the sound and produce a variety of qualities.

10:00—10:15 Break

10:15

5aMU6. Extending the acoustic ensemble through spectral and temporal transformations in real-time. Doug Van Nort (Dept. Architectural Acoust. and Dept. Arts, Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180, vannod2@rpi.edu)

The paradigm of live performance mixing acoustics and electronics has predominantly focused on simple background “tape music,” human players performing highly structured sample-based music (e.g., using the ABELEN LIVE software), or reactive systems that respond to player qualities such as timing, pitch, and so on. In this talk I will present my approach to improvised “laptop performance” that focuses on the transformation of acoustic players in real-time. Rather than simply altering the acoustic content in the manner of an effect processor, the goal is to capture notes and phrases in short-term memory and to re-articulate the material so that it presents a new gestural inflection and timbral content that can be completely novel or suggestive of other players’ sound. The system presented utilizes
a hybrid system combining spectral analysis and feature extraction with block-based temporal processing and a feedback delay network. The interaction paradigm of “scrubbing” the intermediate time/frequency representation is used to generate the ultimate output. The result in an ensemble context is an extended palette that can “keep up” with the musical dialog while eliciting the subtle textural qualities of acoustic players. [This work was supported by NSF Grant 0757454 and CIRMMT/McGill University Fellowships.]

10:35

5aMU7. The digital audio player as aleatory musical instrument. Adam Di Angelo (SIA Acoust., 75 Rockefeller Plaza, 26th Fl., New York, NY 10019, diangelo@gmail.com)

The digital audio file player is an ideal instrument for the performance of aleatory music. The near-instantaneous seek time coupled with random track playback yields the possibility of nonlinear musical arrangements. In this is the potential to create unique and unrepeatable listening experiences—music that constantly renews and reinvents itself, free from the static limitations of the recorded medium. This study examines the history of chance as a compositional and performance tool from 18th century musical dice games to contemporary stochastic methods. It presents examples of works which are designed to be structurally random and outlines the process of employing the digital audio player in the production of variable music. Through such considerations, the role of the conventional mp3 player can expand from passive playback device to become an engine integral to a new kind of listening.

Contributed Paper

10:55

5aMU8. The physics of pitch bending on the vibraphone. Randy Worland (Dept. of Phys., Univ. of Puget Sound, 1500 N. Warner, Tacoma, WA 98416, worland@pugetsound.edu)

The vibraphone is a member of the keyboard (or mallet) percussion family, with tuned aluminum bars typically spanning three chromatic octaves (F3–F6). Extended performance techniques include the use of harmonics and bowing of bars, as well as the pitch bending effect discussed here. After first striking the middle of a bar with a soft mallet in the normal manner, the pitch bend is obtained by pressing a hard mallet onto the bar at a nodal point and then sliding it away from the node as the note sustains. The audible result is a descending pitch, typically of about one semitone. Experimental data on frequency versus location of the hard mallet along the bar are presented. The mass, hardness, and orientation of the sliding mallet relative to the bar are discussed. Interferograms of the lower frequency modes of the bar are also shown. Results are interpreted using a mass perturbation model.

11:10—11:15 Break

11:15—12:00

Participants will give informal demonstrations of several of the techniques discussed.

FRIDAY MORNING, 23 APRIL 2010

GRAND BALLROOM III/IV, 8:00 TO 9:45 A.M.

Session 5aPAa

Physical Acoustics: Theoretical Models

Yves H. Berthelot, Chair

Georgia Inst. of Technology, School of Mechanical Engineering, Atlanta, GA 30332-0405

Contributed Papers

8:00

5aPAa1. Green’s function for the scattered field near the surface of a dissipative elastic medium. Evgenia A. Zabolotskaya, Yuri A. Ilinskii, and Mark F. Hamilton (Appl. Res. Labs., The Univ. of Texas at Austin, Austin, TX 78713-8029)

The problem of scattering in bounded elastic media is relevant to medical ultrasound imaging and detection of buried objects. A method was developed that takes into account the propagation of longitudinal, shear, and Rayleigh waves in an elastic medium. The approach is based on solving the inhomogeneous momentum equation for the Green’s function. The forcing function is modified to accommodate different kinds of excitation that are important for shear wave imaging and seismic wave generation. The solution for the Green’s function is presented as a Fourier integral, and coefficients in the Fourier expansion are calculated by applying boundary conditions on the solid surface and matching conditions in the plane parallel to the surface that contains the source. The method permits the scattered bulk and surface waves to be distinguished from one another. The Green’s function has been presented previously in both the frequency and time domains for a lossless medium. The model is generalized here for elastic media with losses. Besides reducing the amplitude of the response, losses also impair ability to distinguish between the bulk and surface modes of propagation. [Work supported by NIH DK070618.]

8:15


In this paper we will present the analysis of reflected waves of a point source from a two dimensional wedge. Exact method of images is used to calculate the position and strength of image sources. It is shown that the traditional fractional order Bessel function series for the total pressure can be replaced by quadratures of elementary functions.
8:30
5aPAa3. Transition matrix (T-matrix) of circular cylinders embedded in a viscous fluid. Edgar Reyes-Ayon, Daniel Torrent, and Jose Sanchez-Dehesa (Wave Phenomena Group, Polytechnic Univ. of Valencia, Camino de vera s.n., ES-46022 Valencia, Spain)

The present study concerns the scattering of a plane sound wave by a cylindrical scatterer of circular cross section. An infinitely long cylinder embedded in a viscous fluid is here considered. The linearized equations of viscous fluid and the scalar and vector potentials are employed to obtain the T-matrix for two different types of material cylinders: (i) a rigid cylinder and (ii) a fluidlike cylinder with losses. Both cases are also studied and discussed in the limit of long wavelengths. [Work supported by MICIN of Spain and ONR.]

8:45
5aPAa4. A numerical model for acoustic scattering from an inhomogeneous medium. Max Denis, Jing Tsui, Jessica Piper, Kavitha Chandra, and Charles Thompson (Dept. of Elec. and Comput. Eng., Univ. of Massachusetts Lowell, 1 Univ. Ave., Lowell, MA 01854, charles.thompson@uml.edu)

In this work, a numerical method for evaluating the scattering of acoustic waves from an inhomogeneous medium is presented. The internal resonant and multiple scattering effects are modeled by the Padé approximants of a perturbation expansion of the contrast properties between the medium constituents. Due to the local divergence in the Padé approximant solution, a local expansion method about any convergent solution is performed. This allows one to extend the numerical model to higher frequencies. The results show good agreement for geometrically simple scatterers. Based on the simulations, analyses are made of inhomogeneous media micro-structural effects on the scattered field. Of particular interest is the frequency dependence relationship.

9:00
5aPAa5. Single-channel time reversal with Lanczos iterations. Zachary J. Waters (Naval Res. Lab., Code 7130, Washington, DC 20375-5320) and Paul E. Barbone (Dept. of Mech. Eng., Boston Univ., 110 Cummings St., Boston, MA 02215)

Recent investigations of array-based time reversal techniques employing Lanczos iterations have demonstrated convergence properties superior to equivalent methods using power iterations [A. A. Oberai et al., “Efficient time reversal by Lanczos iterations,” J. Acoust. Soc. Am. 123, 3596 (2008)]. In numerical scattering simulations, time reversal techniques developed for a single-channel transducer are applied to study echoes from a thin-walled spherical shell. When power iterations are employed, a narrowband waveform corresponding to the dominant scattering mode of the target is selected. Lanczos iterations converge to this waveform in fewer transmissions and simultaneously identify additional higher-order target scattering features. The performance of both techniques is studied in the presence of varying levels of stochastic and deterministic noise. [Work supported by the Office of Naval Research.]

9:15

One of the limiting factors of performing high-resolution modeling of outdoor sound propagation in the time domain is the treatment of the boundary conditions. Impedance ground boundary conditions have been successfully modeled in the time domain for a variety of porous ground types via the use of recursive convolution methods; however, the applications were for a flat ground. The modeling of sound propagation around buildings (or other obstacles) is usually accomplished by drastically simplifying the building geometry, thereby failing to account for fine-scale, or even some larger-scale, details. Experiments conducted at a fire training facility found that external building features, such as facades and fire escapes, can affect the sound field significantly enough so that standard acoustic diffraction models fail to accurately predict the sound field. Here we develop numerical models that allow us to capture the effects of finer-scale features. Such treatments are necessary for the development of source localization algorithms in urban or other multifarious environments.

9:30
5aPAa7. Energy flux streamlines vs the alternatives for the visualization of energy coupling inside and outside the surface of an ensonified spherical acoustic lens: Preliminary results. Cleon E. Dean (Dept. Phys., P.O.B. 8031, Georgia Southern Univ., Statesboro, GA 30460-8031, cdean@georgiasouthern.edu) and James P. Braselton (Georgia Southern Univ., Statesboro, GA 30460-8093)

Energy flux streamlines are compared with alternative presentations of the energy flux vector field to visualize energy coupling inside and outside the surface of an insonified spherical acoustic lens. The emphasis is on the energy flux streamline as a natural bridge between necessarily approximate ray solutions and full-fledged wave solutions. The present work is the wave solution for the experimental work performed by Kendez C. Parker and Cleon E. Dean. Skills and the interpretation of the energy flux streamline representation due to the spherical geometry of the scatterer are noted.
It was demonstrated recently that air-water interface, which is usually an almost perfect reflector of acoustic waves, becomes anomalously transparent and the power flux in the wave transmitted into air increases dramatically, when a compact sound source in water approaches the interface within a fraction of wavelength [O. A. Godin, Phys. Rev. Lett. 97, 164301 (2006)]. Powerful underwater explosions and certain natural sources, such as underwater landslides, generate very low-frequency waves in water and air, for which both fluid buoyancy and compressibility simultaneously serve as restoring forces. In this paper, analysis of sound transmission through air-water interface is extended to acoustic-gravity waves (AGWs). It is found that, as for sound, the interface becomes anomalously transparent for sufficiently shallow compact sources of AGWs. Depending on the source type, the increase in wave power flux into air due to diffraction effects can reach several orders of magnitude. Physical mechanisms responsible for the anomalous transparency are discussed. Excitation of an interface wave by an underwater source is shown to be an important channel of AGW transmission into atmosphere, which has no counterpart in the case of sound.

The normal reflection and transmission coefficients of a lossy elastic interface, Josif M. Faks (Zel Technologies LLC and NOAA/Earth System Res. Lab., Mail Code R/PSD99, Boulder, CO 80305-3328) and Oleg A. Godin (Univ. of Colorado, Boulder, CO 80305-3328, oleg.godin@noaa.gov)

It was demonstrated recently that air-water interface, which is usually an almost perfect reflector of acoustic waves, becomes anomalously transparent and the power flux in the wave transmitted into air increases dramatically, when a compact sound source in water approaches the interface within a fraction of wavelength [O. A. Godin, Phys. Rev. Lett. 97, 164301 (2006)]. Powerful underwater explosions and certain natural sources, such as underwater landslides, generate very low-frequency waves in water and air, for which both fluid buoyancy and compressibility simultaneously serve as restoring forces. In this paper, analysis of sound transmission through air-water interface is extended to acoustic-gravity waves (AGWs). It is found that, as for sound, the interface becomes anomalously transparent for sufficiently shallow compact sources of AGWs. Depending on the source type, the increase in wave power flux into air due to diffraction effects can reach several orders of magnitude. Physical mechanisms responsible for the anomalous transparency are discussed. Excitation of an interface wave by an underwater source is shown to be an important channel of AGW transmission into atmosphere, which has no counterpart in the case of sound.

The acoustic mirage, like the optical mirage, is a consequence of temperature gradients existing in the air above Earth surface. Here, we report experiments showing that temperature gradients can be used to collimate an airborne ultrasonic beam. A very simple system consisting of two glass tubes heated by an internal resistive coil is used to create a temperature gradient producing the desired effect. Numerical simulations are also performed and results are in agreement with the experimental findings [Work supported by MICINN of Spain.]

The normal reflection and transmission coefficients of a lossy elastic plate are measured and associated according to the outcomes of resonance scattering theory (RST) and S-matrix theory (SMT). The sum and the difference of those coefficients provide valuable tools, the transition conditions, we use to characterize the plate. The absorption in the plate can be expressed via different manners recalled in the following. By means of a dimensionless loss factor, the velocities of the longitudinal and transverse waves of the studied aluminum plate are no more real but complex. The width of a resonance frequency, the transition terms obey the Breit–Wigner resonant form which makes it possible an easy determination of the two parts of the resonant width. There is a good experiment/theory agreement when an imaginary part is added to the Lamé coefficients of the aluminum plate. The loss factor is of the order 5.10–4. On the whole, the effect of the attenuation rises with frequency and reaches a maximal importance at the vicinity of a resonance frequency.

Measurements of sound radiation from binary collisions of polypropylene balls, Joshua Riner and Andi Petculescu (Dept. of Phys., Univ. of Louisiana at Lafayette, P.O. Box 44210, Lafayette, LA 70504, riner@louisiana.edu)

Measurements of sound radiation from binary collisions of polypropylene balls were performed in order to study the collision dynamics and constrain the fraction of incident energy radiated as sound. In the experiments, one ball is released from directly above a stationary target ball. In the first experiment, the incident energy of the incident ball is varied by changing the drop height. The sound is recorded at an angle of 55 deg below the impact plane. Data from this experiment, fitted with a theoretical model, show a possible deviation from Hertzian behavior. The second experiment concerns the radiation pattern of the collision process. The sound pressure was measured as a function of the polar angle \( \theta \). It is assumed that the problem is azimuthally symmetric, i.e., independent of the angle \( \phi \). The results are somewhat surprising: they show a dipole-like pattern that is asymmetric with respect to the impact line. The acoustic energy radiated during the impact, obtained by multiplying the collision time by the sound intensity integrated over a spherical surface centered at the impact point, is estimated as four orders of magnitude smaller than the incident energy. [The work was supported by the Louisiana Board of Regents.]

Silicone oil was found to be fully miscible in “segregated” perfluorinated hydrocarbon (Novoc fluid, by 3M) above a critical temperature slightly above room temperature and at atmospheric pressure. This constitutes a binary fluid mixture, and these have long been used as examples of continuous phase transitions, especially in regard to the phenomenon of critical opalescence [T. Andrews, Philos. Trans. R. Soc. London 159 (1869) and M. Smoluchowski, Ann. Phys. 21 (1906)]. During a phase transition in these mixtures, the fluids become spatially segregated, and the density fluctuation correlations are known to span a large number of length scales [L. Landau and E. M. Lifshitz, Statistical Physics, 3rd ed., Part 1]. This behavior is reflected in massive fluctuations in sound speeds and viscosities in the mixtures near the critical point, and the resulting scattering behavior is probed by actively acoustically interrogating the mixture with techniques including swept-frequency acoustical interferometry and pulse-echo measurements. It is shown that with higher-power acoustical probes, the transition from unary to binary mixture may be prevented or enhanced. [This work was supported by Chevron USA.]
measurements. Formed, which help explain the discrepancy between the CP model and the measurements. Finite-element simulations of both experiments were performed which showed the effects of atmosphere filtering, the thunder waveform described the measured sound speed dispersion approaching the infinite bubble resonance frequency and qualitatively predicted the frequency range of high attenuation, even in this regime of high-void fractions and discrete bubble distribution within the waveguide. Qualitative agreement was found between the CP model and the toroidal bubble measurements. Finite-element simulations of both experiments were performed, which help explain the discrepancy between the CP model and the measurements. [Work supported by Shell.]

A one-dimensional resonator technique was used to investigate the acoustic behavior of water containing large spherical and toroidal encapsulated air bubbles. Tethered balloons and inner tubes, respectively, were used to create the bubbles, which had effective spherical radii of approximately 5 cm. The number of balloons or inner tubes varied from 3 to 6, which resulted in void fractions from 1.5% to 5%. Effective mixture sound speeds for both shapes were inferred from the resonances of a water and bubble-filled waveguide 1.8 m in length. For the spherical bubbles, the Commander and Prosperetti (CP) model [J. Acoust. Soc. Am. 85, 732–746 (1989)] quantitatively described the measured sound speed dispersion approaching the individual bubble resonance frequency and qualitatively predicted the frequency range of high attenuation, even in this regime of high-void fractions and discrete bubble distribution within the waveguide. Qualitative agreement was found between the CP model and the toroidal bubble measurements. Finite-element simulations of both experiments were performed, which help explain the discrepancy between the CP model and the measurements. [Work supported by Shell.]

5aPA6. Acoustic behavior of large encapsulated gas bubbles with resonance frequencies in the 50–100 Hz range. Mark S. Wochner, Kevin T. Hinojosa, Kevin M. Lee (Appl. Res. Labs., The Univ. of Texas at Austin, Austin, TX 78713-8029), Theodore F. Argo, IV, Preston S. Wilson (The Univ. of Texas at Austin, Austin, TX 78712-0292), and Richard S. Mercier (Texas A&M Univ., College Station, TX 77843-3136)

FRIDAY MORNING, 23 APRIL 2010

FRI. AM 11:30

Session 5aPP

Psychological and Physiological Acoustics: Spectral, Temporal, and Complex Auditory Processing

Jennifer Lentz, Cochair
Indiana Univ., Dept. of Speech and Hearing Science, 200 S. Jordon Ave., Bloomington, IN 47405

Yi Shen, Cochair
Indiana Univ., Dept. of Speech and Hearing Science, 200 S. Jordon Ave., Bloomington, IN 47405

Contributed Papers

9:00

5aPP1. Effects of training in increment-detection and intensity-discrimination tasks. Walt Jesteadt, Harisadhan Patra, Melissa Krivohlavek, and Cynthia Rutledge (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131)

Subjects use different cues when asked to detect a brief increment in a longer duration tone (increment detection) than when asked to discriminate between two tones of the same duration that differ only in level (intensity discrimination). To determine whether experience in one task would generalize or interfere with performance in the other, a group of six subjects completed 4000 trials in an increment detection task, followed by 800 trials of intensity discrimination, and final 800 trials of increment detection. A second group of subjects was tested in the reverse pattern, beginning with 4000 trials of intensity discrimination. The increment or signal was a 90-ms, 4-kHz tone. The pedestal or standard was a 490- or 90-ms, 4-kHz tone presented at 70-dB sound pressure level. A two-interval, forced choice adaptive procedure was used to estimate the level of the increment required for 71% correct. The results showed modest improvement over time for two subjects in increment detection and one in intensity discrimination, but at least one subject showed the opposite effect in each task. There was no generalization or interference between the two tasks and no correlation between increment detection and intensity discrimination across the 12 subjects. [Work supported by NIH.]

11:45

5aPP7. The characteristics of thunder on Titan. Andi Petculescu (Dept. of Phys., Univ. of Louisiana at Lafayette, P.O. Box 44210, Lafayette, LA 70504)

Even though Huygens has not detected acoustic signatures of Titanian lightning, electric discharge processes are thought to be very plausible in Titan’s thick atmosphere dominated by nitrogen and methane. When correlated with electromagnetic signatures, thunder will very likely confirm the occurrence of lightning, both during a probe’s descent and/or once it has landed. The presentation has two parts. In the first part, the mechanisms of Titanian thunder generation are discussed, in the context of the most up-to-date predictions of lightning on Titan, as well as relevant Cassini-Huygens data. In the second part, the propagation of thunder in Titan’s troposphere is addressed. Here, the structure, composition, and ambient conditions of the satellite’s atmosphere will be presented and accounted for. Current Titan lightning models predict cloud-to-ground discharge channels of ~20 km. Long-range Titanian thunder simulations are developed based on the far-field superposition of approximately 7000 N-waves arising from 3-m “elementary” sources. Two lightning geometries will be considered: linear and tortuous. (Tortuous channels are synthesized as a Gaussian distribution in the orientations of the 7000 elementary sources.) For illustrative purposes, to show the effects of atmospheric filtering, the thunder waveform described above will also be propagated in Venus’ atmosphere.


Forward masking has been studied extensively, but the mechanism underlying forward masking is still being debated. The aim of this study was to understand the relative contribution of adaptation and temporal integration on forward masking. In the first experiment, gap detection thresholds were measured for symmetrical and asymmetrical noise markers ranging from 10- to 30-dB SL in 5-dB increments. Further threshold measurements were made while keeping the level of the noise marker before the gap at a constant 30-dB SL while varying the level of noise marker after the gap between 10- and 25-dB SL. Noise marker sets with gap detection threshold of 20 ms or less were selected for the second experiment. In the second experiment, participants adjusted gap durations within an asymmetrical noise marker to be perceptually equal to that of symmetrical noise markers. The results will be discussed in the context of a temporal integration model, an adaptation model, and a hybrid model including both temporal integration and adaptation components.
If a noise burst is introduced into the gap between two brief complex tone bursts, pitch discrimination improves to the level observed when the bursts really are continuous. In the experiment described here, it was shown that perceived continuity also affects sensitivity to envelope phase across the gap. Two complex tone bursts with unresolved harmonics and five regular envelope peaks ("pitch pulses") were presented sequentially, separated by a gap of two periods. When the relative phases of the two bursts were varied, such that the inter-pulse interval (IPI) across the gap was varied, the pitch of the whole sequence was little affected, consistent with previous results suggesting that the pitch integration window may be "reset" by a discontinuity. However, when the gap between the two bursts was filled with a noise with the same spectral envelope as the complex, variations in IPI had substantial effects on the pitch of the sequence. It is suggested that the presence of the noise causes the two tone bursts to appear continuous; hence resetting does not occur, and the pitch mechanism is sensitive to the phase discontinuity across the gap.

9:45
5aPP4. Frequency-specific measurements of the temporal modulation transfer function using narrowband noise. Yi Shen and Jennifer Lentz (Dept. of Speech and Hearing Sci., Indiana Univ., 200 S Jordan Ave., Bloomington, IN 47405, shen2@indiana.edu)

Interpretation of frequency-specific measures of temporal acuity is often confounded by the coexistence of spectral and temporal cues. The current study presents a new experimental technique to measure the temporal modulation transfer function (TMTF), in which inherent envelope modulations in narrowband noise are used as the detection target. Modulation depths of the stimuli were manipulated using a signal processing algorithm that introduced no change in the stimuli’s power spectra. In this way, spectral cues cannot cue the presence of modulation. TMTFs were measured using these narrowbands of noise [TMTF-NS (no spectral cues)] at four center frequencies (2, 4, 6, and 8 kHz) and bandwidths of 16, 24, 32, 48, 64, 96, and 128 Hz. TMTFs using low-pass-noise modulated puretones [TMTF-NBM] were also measured for comparison. In cases where the spectral cues are limited in TMTF-NBM (at low-modulation rates or at high-carrier frequencies), thresholds are very similar between the TMTF-NS and the TMTF-NBM, providing validation that this new technique measures temporal acuity similar to standard TMTF methods. In conditions where spectral cues are available in TMTF-NBM (high-modulation rates and lower carrier frequencies), higher thresholds are measured for TMTF-NS and a more reasonable estimate of temporal acuity is obtained.

10:00—10:15 Break

10:15
5aPP5. Another indication that information in “off-frequency” filters enhances sensitivity to envelope-based interaural temporal disparities. Leslie R. Bernstein and Constantine Trahiotis (Depts. of Neurosci. and Surgery (Otolaryngol.), Univ. of Connecticut Health Ctr., 263 Farmington Ave., Farmington, CT 06030, les@neuron.uconn.edu)

A recent study, [Bernstein, L. R., and Trahiotis, C. (2009). J. Acoust. Soc. Am. 125, 3234-3242] employing "raised-sine" stimuli centered at 4 kHz revealed three outcomes concerning discriminability of envelope-based interaural temporal disparities (ITDs). First, graded increases in raised sine exponent led to graded decreases in envelope-based threshold-ITDs. Second, threshold-ITDs decreased with increases in modulation depth. Third, when the modulation depth was 25%, increasing the exponent led to especially large decreases in threshold-ITD. Overall, an interaural correlation-based model including stages mimicking peripheral auditory processing was shown to capture quite well most of those outcomes. Left unexplained, however, was the interaction between modulation depth and raised sine exponent. Here, we report new threshold-ITDs obtained while varying, factorially, depth of modulation (25% or 100%), raised-sine exponent (1.0 or 8.0), and modulation frequency (32, 64, 128, or 256 Hz). Those stimulus parameters were chosen specifically to shed light on the unexplained interaction. A quantitative theoretical analysis of the new data suggests that listeners can enhance their sensitivity to envelope-based ITDs by utilizing "off-frequency" filters centered above 4 kHz. [This research was supported by research grant NIH DC-04147 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.]

10:30
5aPP6. Masker modulation regularity and its effects on comodulation masking release. Emily Buss, Joseph W. Hall, and John H. Grose (UNC School of Medicine, Chapel Hill, NC 27599, ebuss@med.unc.edu)

Several authors have suggested that comodulation masking release is largest when performance in the narrowband masker condition is poor due to perceptual similarities between masker fluctuations and an added signal. The present study tested this hypothesis by manipulating the regularity of masker modulation between but not within listening intervals. The signal was a 500-Hz tone burst, and the masker was a continuous band of noise, either 25- or 1050-Hz wide. After filtering, the masker was amplitude modulated with one of three envelope patterns. The regular envelope was a smoothed 20-Hz square wave, composed of 10-ms raised-cosine ramps and 15-ms steady states. In the irregular envelope conditions, either the duration of the steady state or the modulation depth was randomly selected prior to each modulation period, with one exception: modulation was not randomly perturbed during listening intervals, when the signal might be presented. There were large individual differences, but masker envelope irregularity tended to elevate thresholds, particularly for the 25-Hz masker bandwidth. Results were consistent with more CMR under conditions for which there is greater informational masking in the narrowband masker condition.

10:45
5aPP7. Effects of modulated noise on detection of formant-like stimuli. Peggy B. Nelson (Dept. of Speech-Lang.-Hearing Sci., Univ. of Minnesota, Minneapolis, MN 55455), Magdalena Wojtczak (Univ. of Minnesota, Minneapolis, MN 55455), and Yingjiu Nie (Univ. of Minnesota, Minneapolis, MN 55455)

One problem reported by listeners with hearing impairment (HI) might be that formant transitions in noise are not as salient as for listeners with normal hearing (NH). We hypothesize that a reduced ability to perceive formant transitions in noise could contribute to reduced masking release seen in HI listeners. In this study, we used pure-tone glides and harmonic complex simulations of formant transitions (shifts in intensity increments across frequency range of a formant transition over the transition duration) to test this hypothesis. Pure-tone glides were 30- or 60-ms long and were centered at 1250 or 2500 Hz, frequencies representing second formants in speech. They were presented in isolation or with unmodulated flanking tones (simulating F1 and F3) in quiet, steady noise, and gated noise. NH listeners were only slightly better than HI listeners at detecting the single gliding tone in quiet. Flanking tones adversely affected performance of the HI but not NH listeners. Performance of the NH listeners was somewhat better in gated than in steady noise, whereas the HI listeners performed similarly in steady and gated noise. The role of perception of formant transitions in masking release will be discussed. [Work supported by NIDCD R01-DC008306.]

11:00
5aPP8. The fluctuating-masker benefit for word identification with spectral or fine-structure distortions. Joshua G.W. Bernstein (Army Audiol. and Speech Ctr., Walter Reed Army Med. Ctr., Washington, DC 20307, joshua.g.bernstein@us.army.mil)

Normal-hearing (NH) listeners presented with speech processed to reduce spectral detail or remove temporal fine-structure (TFS) information show a reduced benefit from momentary dips in the level of a fluctuating masker. This has been interpreted as evidence that related deficits may underlie the reduced fluctuating-masker benefit (FMB) observed for the hearing impaired. However, the reduced FMB for processed stimuli could be attributable to the greater signal-to-noise ratio (SNR) required to yield performance levels equivalent to the unprocessed case for unmodulated noise (the reference condition). NH listeners were tested in the identification
of isolated words presented in stationary-noise, interfering-talker, and speech-modulated noise maskers. Different response-set sizes were used to offset performance differences between processed and unprocessed conditions. A validation experiment showed no effect of set size on the FMB when compared at a common stationary-noise SNR. A second experiment compared the FMB for unprocessed, spectrally smeared, and noise vocoded speech. When compared at the same SNR and percent-correct level (but different set sizes), processed and unprocessed stimuli yielded similar FMB. This suggests that for the conditions tested here, spectral or TFS distortions do not directly impair the ability to listen in the gaps of a fluctuating masker. [Sponsored by the Oticon Foundation.]

11:15
5aPP9. Psychometric properties of the coordinate response measure corpus with various types of background interference. David A. Eddins (Dept. of Otolaryngol., Univ. of Rochester, Rochester, NY 14642 and the Int. Ctr. for Hearing and Speech Res., Rochester Inst. of Technol., Rochester, NY 14623 USA.) and Chang Liu (Univ. of Texas at Austin, Austin, TX 78712)

The prevailing view is that newborn phonetic perception is tabula rasa because of poor transmission of the acoustic features of phonemes to the fetus. However, vowel information may be at least intermittently clear in utero. We tested 80 neonates (M = 32.8 h old, range 7–75) in the US and Sweden with English and Swedish vowels using an infant-controlled sucking procedure. Sucking activated 17 stimuli (a prototype and 16 variants) from the same vowel category, either the English /i/ or the Swedish /y/. Infants sampled through all 17 stimuli, presented randomly, one time. The dependent measure was mean number of sucks per stimulus. Results showed that the Foreign Vowel Group had significantly greater means to the prototype than the Native Group. A within-group analysis showed another difference. Infants in the Foreign Group had significantly more sucks to the prototype compared to the variants, whereas the Native Group treated the prototype and variants equivalently. These results require us to re-evaluate assumptions about availability of speech sounds to the fetus and the state of speech perception at birth.

5aSC2. Longitudinal analysis of vowels in infant-directed speech. Kerry E. McColgan, Nan Bernstein Ratner, and Rochelle Newman (Dept. of Hearing and Speech Sci., Univ. of Maryland, 0100 Lefrak Hall, College Park, MD 20742, kmccolgan@hesp.umd.edu)

Adult-directed conversational speech (ADS) is characterized by a high level of acoustic imprecision which should theoretically cause problems for infant language learning. Whether infant-directed speech (IDS) is clearer than ADS is disputed. However, differences in prosody [Fernald (1989)], utterance length [Cooper (1997)], and acoustic features [Bernstein Ratner (1984) and Malsheen (1980)] have been found. Such differences have the potential to assist children in decoding language input [Fernald (1985)]. In prior research, the child’s age and language ability appear to affect speech clarification [Bernstein Ratner (1996)]. This study examines whether vowel clarification is a function of the infants’ age. More than 75 mother-infant dyads are being followed from 7 months to 2 years of age. Data reported here are from 10 dyads whose data have been fully analyzed. Plots of the first and second formant values for vowels in words spoken in IDS and ADS were compared at 7 and 11 months of age. Vowels to 7-month-old infants appear to show less clarification than those to 11-month infants, although individual patterns of clarification are evident. This study analyzes whether such variability in vowel clarification affects the rate of early infant vocalizations prior to language production. [Work supported by NSF BCS074512.]
within categories over time to give an understanding of the development of children's vowel space. Changes in children's ability to differentiate among vowels in F1-F2 space over time are investigated. The development of vowel space is considered in the context of growth of and changes in the proportions of the vocal tract and improved control of the articulators. [Work supported by NIDCD-0001247 to CReSS LLC.]

5aSC4. Neural coding of formant-exaggerated speech in infants and adults. Yang Zhang, Sharon Miller, Tess Koerner, and Edward Carney (Dept. of Speech Lang.-Hearing Sci., Univ. of Minnesota, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

Speech scientists have long proposed that formant-exaggerated speech plays an important role in phonetic learning and language acquisition. However, there have been very little neurophysiological data on how the infant brain and adult brain respond to formant exaggeration in speech. We employed event-related potentials (ERPs) to investigate neural coding of formant-exaggerated speech sounds. Two synthetic /i/ vowels were modeled after infant-directed speech data and presented in alternating blocks to test the effects of formant exaggeration. The fundamental frequencies of the two sounds were kept identical to avoid interference from exaggerated pitch level and range. For adult subjects, non-speech homologs were also created by using the center frequencies of the formants to additionally test whether the effects were speech-specific. In the infants (6 to 12-month olds), ERP waveforms showed significantly enhanced N250 and sustaining negativity responses for processing formant-exaggerated speech. In adults, enhancement was observed in the N100 component for the speech stimuli but not the homologous non-speech sounds. Collectively, these results provide the first evidence that formant expansion in infant-directed speech enhances neural activities for phonetic encoding, which may facilitate phonetic learning and language acquisition regardless of the age factor [Zhang et al. (2009). Neuroimage 46, 226–240].


Past research on children with cochlear implants (CIs) focused on improvements in speech perception, but few speech production studies have been reported. The literature on speech development in prelingual children with CIs is particularly sparse. This is partially due to difficulty transcribing young children’s productions. The aim of this study was to examine whether developmental changes in vowels can be captured without phonetically transcribed data. The subject, a prelingually deaf child 22 months old at the start of the study, is one of six children with CIs participating in an ongoing longitudinal study. Audio and video data were collected once or twice per week during auditory/verbal therapy at the Alfred I. du Pont Hospital for Children. One author perceptually categorized the child's vowel-like vocalizations in terms of broad height (HIGH, MID, and LOW) and place (FRONT, MID, and BACK) features. Acoustic features (barkscale cepstral coefficients) were then used to train linear discriminant functions on the same vowel data and used to predict perceptual vowel categories. In preliminary results 62% of the tokens were correctly classified for place and 67% for height. Results on generalizability of the features will be presented. [Work supported by NIDRR Grant H133E080003 to the RERC on Hearing Enhancement.]

5aSC6. Sensorimotor maps and vowel development in English, Greek, and Korean: A cross-linguistic perceptual categorization study. Benjamin Munson (Dept. of Speech-Lang.-Hearing Sci., Univ. of Minnesota, 164 Pillsbury Dr., Minneapolis, MN 55455, muns0005@umn.edu), Lucie Mnard (Univ. des Quebes, Montreal, QC C5 3P5, Canada), Mary E. Beckman (Ohio State Univ., Columbus, OH 43210), Jan Edwards, and Hyunju Chung (Univ. of Wisconsin, Madison, WI 53705)

Learning to speak involves learning the association between different articulatory maneuvers and their associated auditory-perceptual characteristics. These associations may change over the course of development as a consequence of the nonlinearities inherent in vocal-tract growth. A set of recent studies by Menard and colleagues has examined these developmental changes by studying adults' categorizations of synthesized vowels based on an articulatory model of the growing vocal tract. These have shown that vowels modeled on younger vocal tracts tend to be perceived as more front than those modeled on older vocal tracts. They have also shown language-specificity in the vowels that are identified in growing vocal tracts, with French listeners labeling fewer sounds as /u/ than English ones, presumably because the high-vowel space of French includes three vowels /i y u/, while English has only two. The current experiment expands on this by examining the perception of synthesized vowels based on seven vocal-tract growth stages by speakers of three languages with very different vowel systems. American English, Korean (seven vowels), and modern Greek (five vowels). Preliminary analyses show language-specific patterns in the perception of vowels in developing vocal tracts. [Work supported by NIDCD 02932 and NSF grants BCS0729140 and BCS0729277]

5aSC7. Korean listeners' sensitivity to language-specific phonetic details of children and adults' vowel production in five different languages. Hyunju Chung, Jan Edwards, Gary Weismer, Eunjong Kong (Dept. of Communicative Disorder., Univ. of Wisconsin-Madison, 1500 Highland Ave., Madison, WI 53705, hchung23@wisc.edu), Benjamin Munson (Univ. of Minnesota, Minneapolis, MN 55455), and Mary E. Beckman (Ohio State Univ., Columbus, OH 43210-1298)

Systematic differences in vowel production across languages have been demonstrated, even for so-called “shared” vowels. Starting from 10 months of age, language-specific vowel production patterns emerge in babbling [Boysson-Bardies et al. (1989)]. Vowel production data from a word repetition task in our laboratory also showed language-specific patterns emerging at 2 years of age. While there is research on adult and infant perception of non-native vowels produced by adults, there is little work on adult perception of children’s non-native vowels of different languages. The current study examined Korean adult listeners’ perception of vowels from five languages: Mandarin Chinese, Cantonese, English, Japanese, and Korean modeled by 2-, 5-year-old children, and adults. Consonant-vowel and vowel stimuli were extracted from our vowel production data. Target vowels included /i/, /a/, and /u/, following velar and alveolar stops. Twenty native Korean adult listeners were asked to identify the vowels and rate their goodness. Results will be analyzed using both the formant frequencies and listeners’ judgments using a mixed effect model. [Supported by a Fulbright fellowship, NIDCD Grant 02932, and NSF Grant 0729140.]


Investigations of auditory feedback control of supra-glottal articulatory movements traditionally have focused on quasi-static speech sounds like monophthongs. Only recently has attention been directed toward time-varying sounds. To study the role of auditory feedback in the formation of formant trajectories in time-varying vowels, we perturbed the second formant frequency (F2) trajectory as 20 native speakers produce the Mandarin triphthong /iau/ and measured their patterns of auditory-motor adaptation. The subjects changed F2 in directions opposite to the perturbations, not unlike adaptive responses seen in previous studies. Surprisingly, concurrent with the F2 compensations, the subjects significantly altered the first formant (F1), the unperturbed formant, in ways that assisted cancellation of the perturbations to the location and orientation of the formant trajectory in the F1×F2 space. Care was taken to rule out that the F1 changes were byproducts of the F2 compensations. These observations indicate that F1 and F2 are planned in a coordinated fashion for the triphthong /iau/, a property of the speech system not yet revealed through studying quasi-static sounds. The implication of this F1-F2 synergy for the spatiotemporal nature of vowel auditory targets will be discussed. The patterns of generalization to a set of unperturbed vowels will also be presented.

5aSC9. Position and height asymmetries in hiatus resolution: Are they phonetically driven phenomena? Hijo Kang (Dept. of Linguist., Stony Brook Univ., Stony Brook, NY 11794-4376, hikang@ic.sunysb.edu)

Hiatus is not preferred in many languages and is resolved in various ways. Two typological patterns have been found in hiatus resolution. When hiatus is resolved by a weakening (deletion or gliding) of one vowel, (1) V1 is more likely to be the target than V2 [Casali (1996)] and (2) high vowels are more likely to be the target than non-high vowels [Rosenhall (1997)]. If
these patterns are due to human articulatory and/or auditory mechanisms [Ohala (1993)], the patterns would be expected to reflect variations in ordinary speech. In this study Korean CV1V2 nonce words produced by six Korean speakers are analyzed to test whether V1 and high vowels in a realized hiatus are actually weakened in terms of duration. The results showed that V1 is shorter than V2 irrespective of speech rate [F(1.5) = 12.90, P < 0.02], which was not found in CVVC or CVGV words. It was also found that high vowels are shortened in fast speech to a more degree than non-high vowels [F(1.5) = 7.16, P < 0.05]. The formant analysis showed that vowel-to-vowel coarticulation effects are greater in CV1V2 than in CVVC, which means that vowels in hiatus are more vulnerable to misperception and consequently, sound change.

5aSC10. Lexical frequency and Japanese vowel devoicing. Kuniko Nielsen (Dept. Linguist., Oakland Univ., 320 O’Dowd Hall, Rochester, MI 48309, nielsen@oakland.edu)

The current study aims to examine the relationship between Japanese vowel devoicing and lexical frequency in light of the probabilistic reduction hypothesis [Jurafsky et al. (2001)], which predicts that more probable words are (phonetically) reduced more than less probable words. In Tokyo Japanese, high vowels become devoiced when they occur between two voiceless consonants. Similar devoicing processes occur in many languages and are generally considered to be part of the vowel reduction processes [e.g., Dauer (1980); Kohler (1990)]. Although Japanese vowel devoicing has been studied extensively from both phonetic and phonological viewpoints, its lexical property, namely, the effect of lexical frequency, is as of yet unknown. If Japanese devoicing is a phonetic reduction process, we would expect to see a stronger degree of devoicing among words with higher lexical frequency. To evaluate the frequency effects of Tokyo Japanese high vowels, we recorded 150 words which include various devoicing environments. Contrary to the prediction by the probabilistic reduction hypothesis, our preliminary results revealed a significantly higher rate of devoicing [F(1,22) = 24.89, P < 0.01] as well as shorter vowel duration [F(1,22) = 4.61, P < 0.05] among low-frequency words than high-frequency words. The implications of these results on the nature of Japanese vowel devoicing will be discussed.

5aSC11. Comparison of the formant frequencies F3 and F4 on a three-dimensional vowel chart. Toshio Isei-Jaakola (Dept. of English Lang. & Culture, Chubu Univ., 1200 Matsumoto, Kasakagi, Aichi 487-8501, Japan, tiseij@isc.chubu.ac.jp); Takatoki Naka (Chukyo Univ., Toyota, Aichi 470-0393, Japan), and Keikichi Hirose (Univ. of Tokyo, Bunkyo, Tokyo 113-8566, Japan)

We have developed a three-dimensional vowel chart that should fit into a quadrilateral IPA vowel chart better and later we have added more visually effective functions to it. On the chart we can see not only the locations of F1 and F2 with two scales of x and y, where x is F1 and y F2, but also F1, F2, and F3 as well, with three scales of x, y, and z, where x is F1, y F2, and z F3. Thus, this enables us to compare the vowel locations of any language (multilingual vowel presentations), vowels uttered by male with those of female (thus, gender difference), vowels of adults with those of children (thus, age difference), foreign language learners with target vowels, or vowels uttered by the same language learners. This time, we focus more on the roles of the respective formants, particularly on the roles of F3 and F4. It has been clarified that F3 is more related to lip-spread. We want to show that F4 may be more related to lip-protrusion. For this purpose, we have further added a selection of F3 and F4 functions to the chart. [Work supported by JSPS and Chubu University Grant (A.1.)]

5aSC12. Psychometric functions of vowel detection and vowel identification in multi-talker babble. Maggie Miller, Jessica Walker, Kayla O’Brien, and Chang Liu (Dept. of Commun. Sci. and Disorder, The Univ. of Texas at Austin, 1 Univ. Station A1100, Austin, TX 78712)

Psychometric functions of vowel detection and vowel identification were measured in multi-talker babble for young normal-hearing listeners. A four-interval forced-choice procedure was used to examine the accuracy of vowel detection in babble with speech level presented from 0- to +15-dB sensation level relative to vowel detection thresholds obtained with method of limits. The accuracy of vowel detection was significantly influenced by vowel category and sensation level. The threshold of vowel detection for each vowel and each listener was defined as the speech level at which 71% accuracy of vowel detection was reached. Vowel detection was then measured in babble with vowel levels presented from 0- to 12-dB sensation level relative to individual thresholds of vowel detection, using a close-set 12-choice procedure. Results suggest that vowel identification was significantly affected by vowel category and sensation level. Altogether, the results of vowel detection and vowel identification indicate that, given the same signal-to-noise ratio, vowels are not equally audible and identifiable, possibly due to the fact that some vowels are more audible than others, and that the slope of psychometric functions of vowel identification is vowel-dependent.

5aSC13. Vowel-inherent spectral change in isolated vowels and consonant vowel consonants. Peter F. Assmann (School of Behavioral and Brain Sci., The Univ. of Texas at Dallas, Richardson, TX 75083), Terrance M. Nearey (Univ. of Alberta, Edmonton, AB T6G 2E7, Canada), and Michael Kiefte (Dalhousie Univ., Halifax, NS B3H 1R2, Canada)

To study the interaction of vowel inherent spectral change (VISC) and consonant context on vowel formant patterns, we recorded 15 vowels /i, æ, a, o, ɔ, u, ɤ, ø/ in each of 14 consonant environments (pVp, pVb, bVp, bVb, pVt, T, Vd, dVt, dVd, kVk, gVk, gVg, hVd, and V) spoken by 10 men and 10 women from the north Texas region. Analysis of vowel formant frequency trajectories confirmed the reliable presence of VISC across talkers for a range of consonant environments. Vowels showing stable patterns of VISC included those acknowledged as diphthongs in most North American English dialects, /æt/ [æt], /æt/, /æt/, /æt/, as well as /ɪt/, with more variable movement patterns for /ɪt/, /ɪt/, and /ɪt/. Of particular interest were cases where the formants F1 and/or F2 showed a “switchback” pattern of movement, with initial movement in the expected direction for the vowel (i.e., F1/F2: Euclidean distance between PSP and PBP) to subsequently switch toward the “locus” for the final consonant. Parallel analyses of VISC in three dialects (north Texas, Nova Scotia, and Alberta) are currently underway and will be reported at the meeting along with statistical modeling results [Nearey (this meeting)].

5aSC14. Assessing acoustic measures of the spontaneous phonetic imitation of vowels. Molly E. Babel (Dept. Linguist., Univ. of British Columbia, 2613 West Mall, Vancouver, BC V6T 1Z4, Canada, mibel@interchange.ubc.ca) and Benjamin Munson (Univ. of Minnesota, Minneapolis, MN 55455)

It is well established that people imitate fine phonetic detail of another talker in shadowing tasks [Goldinger, Psych. Rev. 105, 251–279 (1998)] and in interactive conversation [Pardo, J. Acoust. Soc. Am. 119, 2382–2393 (2006)]. That is, the acoustic characteristics of a model talker’s production (MTP) are more similar to a participant’s shadowed production (PSP) than they are to that participant’s baseline production (PPB), typically elicited in a reading task. This presentation compares two methods of assessing the acoustic distance between the vowels in PSP and PPB monosyllabic words taken from two previous studies [Babel, thesis, University of California (2009); Kaiser & Munson, unpublished]. One of these measures is the F1/F2: Euclidean distance between PSP and PPB vowels. This measure does not take into account the direction of the difference. The other [adapted from Titze, Principles of Voice Production (1994)] compares the distance and direction of the F1 and F2 values in the PSP and PPB vowels relative to those of the MTP. The merits of each of these methods are assessed by comparing them to measures of listener judgments of the similarity of PSPs and PPBs to the MTPs in AXB perception tasks.

5aSC15. The relationship between fundamental frequency and vowel quality. Santiago Barrera and Terrance M. Nearey (Dept. of Linguist., Univ. of Alberta, 4-32 Assiniboia Hall, Edmonton, AB T6G 2E7, Canada, sbarrera@ualberta.ca)

There is disagreement over the role fundamental frequency (F0) plays in the determination of vowel quality. Some claim [e.g., D. Smith et al., J. Acoust. Soc. Am. 117, 305–318 (2005)] that changes in F0 do not affect vowel quality at all. Others claim there is relationship between F0 and vowel quality, whether direct or indirect [see T.M. Nearey and P. F. Assmann, in Experimental Approaches to Phonology, edited by M. J. Solé and P. S. Beddor (Oxford University Press, Oxford, 2007), pp. 246–269 for a review]. To test these theories, a series of experiments was carried out in which participants were asked to make simultaneous speaker and vowel quality judgments. Participants were presented with a synthetic vowel continuum
spanning from [æ] to [ʌ] matched with several different f0s and higher formants. Participants were asked to rate vowel quality on a continuous scale ranging from completely [æ] to completely [ʌ] and to indicate the size and gender of the speaker. Results indicate that there is a complicated, indirect relationship between f0 and vowel quality, and that shifts in vowel quality caused by changes in f0 are a result of changes in assumed speaker properties.

5aSC16. Improved representation of variance in measures of vowel merger. Lauren Hall-Lew (Phonet. Lab., 41 Wellington Square, Oxford OX1 2JF, UK)

Current measures of vowel merger, such as the Euclidean distance between average vowels, have only been able to capture some of the variability between two given vowel clusters. Reliance on averages obscures the amount of variability within a given vowel class, while other techniques, such as calculating distance between minimal pairs, rely on few tokens per speaker. Hay et al. (2006) introduced an alternative approach that accounts for the variability between two vowel clusters, taking formant values as input, rather than averages. The measure is the Pillai-Bartlett statistic [Baayen (2008)], an output of a multivariate analysis of variance (MANOVA), which represents the proportion of one variance that can be predicted by another variance. A higher Pillai value indicates a lower degree of overlap between two vowel clusters in F1/F2 space. Since the value is derived from a MANOVA, Pillais can account for known internal factors influencing the production of merger, such as phonological environment, thereby reducing the need to obtain minimal pair lists. This talk argues for using Pillais as measures of merger by comparing results from low back merger in Californian English [Hall-Lew (2009)] with the analysis of front vowel merger by Hay et al. (2006).

5aSC17. An acoustic study of front rounded vowels in Shetland dialect. Man Gao and Peter Sundkvist (Acad. of Humanities and Media Studies, Dalarna Univ., SE-79160 Falun, Sweden)

This paper presents an acoustic analysis of front rounded vowels (FRVs) in the dialect spoken in the Shetland Islands, the northernmost locality of the British Isles. FRVs are typologically marked and estimated to occur in only 6.6% of the world’s languages [I. Maddieson, in Haspelmath et al. The World Atlas of Language Structures (2005)]. Their occurrence in the Shetland dialect is, at least partly, attributable to a Scandinavian substratum language. There is significant variation across the archipelago regarding several aspects such as (1) the number of lexically contrastive FRVs, (2) phonetic quality (close to half-close), (3) contrastive length, and (4) lexical distribution and support. This paper presents an investigation of three speakers from one locality in which FRVs have retained firm lexical support. The issues addressed concern the dialect’s overall acoustic vowel space (based on F1, F2, and F3), the position of FRVs within the acoustic space, and what the contrasts among FRVs and other adjacent vowels appear to rest on acoustically. Special focus is directed to phonetic contexts that support the greatest number of vowel contrasts and display the most crowded acoustic vowels spaces.

5aSC18. A new non-linear regression model for formant trajectories in English monosyllables incorporating dual targets for vowels. Terrance M. Nearey (Dept. of Linguist., Univ. of Alberta, 4-32 Assiniboia Hall, Edmon ton, AB, T6G 2E7, Canada, t.nearey@ualberta.ca)

A new non-linear regression model is proposed to characterize the formant trajectories of the vocalic portion of English CV/CVC syllables in the data described by Hillenbrand et al. [J. Acoust. Soc. Am. 109, 748–763 (2001)]. The modeling framework builds on work of Broad and Clermont [J. Acoust. Soc. Am. 81, 155–165 (1987)], wherein formant trajectories were modeled via three additive components: (1) a single vowel target, (2) an exponential approach (in time from onset) toward the vowel target from an initial consonant onset value, and (3) an exponential approach (in time from offset) toward the vowel target from a final consonant offset value. The new model extends this to allow for a dual specification (nucleus + offglide) of the vowel targets [T. Nearey and P. Assmann, J. Acoust. Soc. Am. 80, 1297–1308 (1986)]. Initial results suggests that inclusion of a second vowel target provides substantial reduction (about 40%, 23% and 7% respectively for F1, F2, and F3) of error variance on trajectories averaged across 12 speakers. More detailed statistical analyses of variants of the new model are underway and will be reported for both the data described above and that reported on by Assmann et al. [this meeting].

5aSC19. The effect of increased loudness on anticipatory coarticulation of steady-state formants in a vowel-consonant-vowel context for patients with Parkinson’s disease. Jennifer A. Kamphaus, Joan E. Sussman, Elaine T. Stathopoulos, Jessica E. Huber, Kelly C. Richardson, and Brit A. Boyarsky (Dept. of Communicative Disord. and Sci., Univ. at Buffalo, SUNY, 3435 Main St., Buffalo, NY 14214)

The current investigation examines the effect of increased loudness on anticipatory coarticulation of the vowels [ɪ, ʌ, ʊ] on the preceding schwa vowel for patients with Parkinson’s disease. Four speakers, three males and one female with speech severity levels between 8% and 49%, are involved in pre-treatment, treatment, and post-treatment conditions. During treatment, participants use an auditory device which plays speech-babble into one ear while they are speaking. The speech-babble competition results in an automatic loudness increase (Lombard effect) in the talker wearing the device. Participants produced target nonsense CVC words embedded in carrier sentences at comfortable loudness levels and at an average of 3 dB above that level. The steady-state F1, F2, and F3 formant frequencies of the preceding [a] were analyzed to determine whether increased loudness results in increased anticipatory coarticulation from the [ɪ, ʌ, ʊ] vowels in the CVC words, as treatment progresses. Results using steady-state formant frequency measures from this study will be compared with previous studies that used F2 trajectory measures [Tjaden and Sussman (2006); Weismer et al. (1995); Tjaden (2003)]. (Work supported by NIH IR01DC009409-01)

5aSC20. Investigating within and between talker variability in the optical characteristics of American English vowels. Edward T. Auer, Jr. and Laura Welch (Dept. of Speech-Lang.-Hearing, Univ. of Kansas, 1000 Sunnyside Ave., Rm. 3001, Lawrence, KS 66045,auer@ku.edu)

Previous research investigating vowel production provides evidence of substantial within and cross-talker variation in acoustic characteristics. The current study is designed to investigate whether similar within and cross-talker variation exists in the optical characteristics of American English vowels. Additionally, potential shared acoustic-optical variation across tokens will also be examined. Measures taken from three-dimensional motion data for a set of 13 markers positioned around the lips, cheeks, and chin sampled at 100 frames/s with millimeter spatial resolution along with simultaneous video (50 frames/s) and audio for each of ten repetitions of 11 vowel spoken by multiple talkers will be reported. Preliminary analysis of motion and acoustic data from 12 repetitions of 11 vowels spoken by one male talker provides evidence of substantial within vowel variation in both vertical and horizontal lip separations measured at the midpoint of the vowel as determined by analysis of the acoustic signal. Correlations were observed between horizontal lip separation and second formant frequency (r = 0.48) and third formant frequency (r = 0.66) at the vowel midpoint. No correlations were obtained with vertical lip separation measure. Results will be discussed in terms of potential implications for talker intelligibility in visual only presentation conditions.

5aSC21. Effects of ambient pressure and gas mixture on a numerical model of subglottal acoustics and vocal spectra. Steven M. Lulich (Dept. of Psych., Washington Univ., 1 Brookings Dr., St. Louis, MO 63130, sulich@artsci.wustl.edu)

Divers’ speech in hyperbaric helium environments becomes significantly distorted, making communication between divers and with surface-based personnel difficult. Previous research identified two basic mechanisms underlying this distortion. First, the speed of sound increases significantly in helium, causing a linear increase in the second and higher formant frequencies. Second, in hyperbaric conditions the specific impedance of air in the vocal tract approaches that of the vocal tract walls, leading to a nonlinear increase in the first formant frequency. Effects of ambient pressure and gas mixture on subglottal acoustics and subglottal-vocal tract coupling in vowel spectra have not been sufficiently investigated. In this paper, such an investigation is carried out using a numerical model of subglottal and vocal acoustics. It is shown that subglottal resonance frequencies increase linearly as the speed of sound increases, and the corresponding subglottal pole-zero pairs in vowel spectra become disproportionately prominent as
ambient pressure increases. Increased acoustic coupling between the subglottal system and the vocal tract in hyperbaric helium therefore constitutes a third significant source of speech distortion. The results have implications for understanding subglottal coupling in vowels and for heliox speech unscrambling techniques. [This work was supported in part by NSF Grant No. 0905250.]

5aSC22. Developing vowel mappings for an interactive voice synthesis system controlled by hand motions. Karl I. Nordstrom (Media and Graphics Interdisciplinary Ctr., Univ. of British Columbia, Forest Sci. Ctr. Bldg., FSC 3640-2424 Main Mall, Vancouver, BC, V6T 1Z4, Canada, karl@karlnordstrom.ca), Sidney Fels, Cameron D. Hassall, and Bob Pritchard (Univ. of British Columbia, Vancouver, BC, V6T 2Z2, Canada)

This study investigates vowel mappings for a voice synthesizer controlled by hand gestures for artistic performance. The vowel targets are on a horizontal plane navigated by the movement of the right hand in front of the performer. Two vowel mappings were explored. In one mapping, the vowels were evenly distributed in a circle to make the vowel targets easier for the performer to find. In the other mapping, the vowels were arranged according to the F2 versus F1 space. Linear hand motions were then made through the vowel space while plotting the formant trajectories. The evenly distributed mapping resulted in formant trajectories that were not monotonous; the F1 and F2 pitch contours varied up and down as the hand carried out the linear motions. This had the unintended result of producing multiple diphthongs. In contrast, the F2 versus F1 mapping enabled the performer to create monotonous formant trajectories and the perception of a single diphthong. The performer found it easier to speak and sing through the system when a single linear hand motion resulted in a single diphthong. [This project was supported by Canada Council for the Arts, Natural Sciences and Engineering Council of Canada, and Media and Graphics Interdisciplinary Centre.]

5aSC23. Psychometric functions for rough voice quality. David A. Eddins (Univ. of Rochester, 2365 S. Clinton Ave., Ste. 200, Rochester, NY 14618, david_eddins@urmc.rochester.edu) and Rahul Shrivastav (Univ. of Florida, Gainesville, FL 32611)

Dysphonic voices arising from laryngeal lesions are generally believed to vary across three perceptual dimensions—breathiness, roughness, and strain. This experiment sought to evaluate the psychometric functions for roughness and to determine the acoustic cues for its perception. Ten normal-hearing listeners, all native speakers of American English, were recruited for this experiment. The perception of roughness in dysphonic voices was evaluated for 34 voices from the Kay Elemetrics Disordered Voice Database (Kay-Pentax, Inc., Lincoln, NJ). Listeners compared the roughness of these test stimuli against that for a sawtooth wave with a 40-Hz sawtooth-square-amplitude modulation (AM). The modulation depth of the sawtooth wave varied from low to high in a paired comparison task. Psychometric functions for roughness were derived by plotting the listener response for each stimulus pair against the modulation depth of AM sawtooth wave. The data for each stimulus were fitted with a logistic function, and various parameters (threshold, slope, and saturation point) were computed. These parameters were then compared to various acoustic measures obtained from each voice to determine candidate acoustic cues for roughness.

5aSC24. Glottal aperture monitoring with external lighting and sensing photoglottography. Kiyoshi Honda (LPP, UMR7018 CNRS-Univ. Paris 3, 19 rue des Bernardins, 75005 Paris, France), Shinji Maeda (TELecom ParisTech, F-75634 Paris Cedex 13, France), and Tatsuya Kitamura (Koan Univ., Kobe 658-8501, Japan)

A non-invasive photoglottographic method has been developed for monitoring glottal aperture changes during speech for the purpose of phonetic and clinical studies. The system includes light-source and sensor units both placed externally on the neck. An LED light source on the side of the neck illuminates the hypopharynx diffusely, and a photo-sensor unit on the front neck below the cricoid cartilage detects light passed through the glottis. An ambient light rejection circuit was newly added to avoid the effect of room light. The photoglottography (ePGG) system is free from interference due to tongue retraction and thus operational both in high- and low-vowel environments, while it is susceptible to articulatory movements of the jaw and larynx. We will present new ePGG/airflow data to explain why the apparent word-initial strengthening of glottal opening occurs in our ePGG, as often observed in other previous PGG data. [Work supported by EASPV/ANR and Kakenhi 2100071.]

5aSC25. Breathy phonation in Gujarati: An acoustic and electroglostographic study. Sameer ud Dowla Khan (Dept. of Linguist., Cornell Univ., 203 Morrill Hall, Ithaca, NY 14853, sameerudowlakhan@gmail.com)

Recent electroglostographic research on Gujarati [Khan (2009)] reveals that speakers consistently distinguish breathy and modal vowels (e.g., /ba/ “twelve” vs /baR/ “outside”; capitalization represents breathiness) by both closed quotient and closing velocity. Despite this interspeaker uniformity measured at the voice source, the acoustic correlates of the contrast measured in the speech output vary considerably across speakers. While some speakers mark breathiness with a larger H1-H2 difference, others use lower periodicity (CP) or greater changes in intensity (rms energy). It is unclear how breathy vowels are contrasted with modal vowels following breathy consonants (e.g., /baR/ “outside” vs /baR/ “burden”) or to what extent the preceding consonant manner can affect the spectral acoustics of the vowel (e.g., /paR/ “mountain” vs /haR/ “vehicle”). To better capture the variation in both vowel and consonant phonations, the current study examines acoustic and electroglostographic data collected from naturalistic productions of near-minimal sets from a larger pool of Gujarati speakers, using a wider range of vowel and consonant types. Preliminary results suggest that while all speakers distinguish phonation types in both consonants and vowels based on electroglostographic measures, the acoustic cues in the output are far more complex. [Work supported by NSF.]

5aSC26. Effects of asymmetric vocal fold activation on phonation. Dinesh K. Chhetri and Juergen Neubauer (Div. of Head and Neck Surgery, 62-132 CHS, 10833 Le Conte Ave., Los Angeles, CA 90095, dchhetri@mednet.ucla.edu)

The effects of asymmetric vocal fold activation on phonation were studied using an in vivo canine model and graded neuromuscular stimulation of the recurrent laryngeal nerve (RLN). One RLN was stimulated at 21 levels of graded stimulation, from threshold activity to maximal contraction, at three levels of contralateral vocal fold contraction (low, medium, and high). Phonation onset pressure (Pth) was recorded at each condition. The effects of airflow and cricoarytenoid (CT) muscle activation were studied by varying airflow and CT activation levels. Pth remained relatively constant over nearly all conditions while the level of graded stimulation needed to reach Pth varied. Increasing airflow and higher level of contralateral vocal fold activation allowed Pth to be reached at lower levels of graded stimulation, while activation of CT muscles had the opposite effect. At maximal CT activation, Pth was reached only at high levels of contralateral vocal fold activation but Pth was doubled. Results show that in vocal fold paresis /paralysis, phonation onset can be achieved by increasing either airflow or activation level of the normal vocal fold, and that CT activation increases the phonatory effort. These results provide further insights into the compensatory mechanisms required for phonation in these common laryngeal pathologies.

5aSC27. An acoustic analysis of pulmonic ingressive speech in the Shetland Islands. Peter Sundkvist, Man Gao (Acad. of Humanities and Media Studies, Dalarna Univ., SE-79160 Falun, Sweden, psn@du.se), and Gunnar Melchers (Stockholm Univ., SE-10691 Stockholm, Sweden)

The use of a pulmonic ingressive airstream mechanism in the pronunciation of certain discourse particles, typically variants of “yes” and “no,” is a well-known and salient feature of Scandinavian languages. It has been suggested, however, that this may be a more general North Atlantic phenomenon—occurring as far west as Newfoundland and New England—which spread via migration and trade routes. Unfortunately, there seem to be very few audio recordings available from areas other than Scandinavia and Newfoundland (perhaps partly attributable to various elicitation difficulties) and very little acoustic analysis has been presented [E. Thom, MA thesis, UCL (2005); R. Eklund, J. Int. Phonetic Assoc. 38, 235–252 (2008)]. This paper contains a study of ingressive discourse particles in the Shetland Islands, which have strong historical links to Scandinavia. A significant number of ingressives were found in field recordings from 1980–1982. This paper presents an acoustic pilot study of ingressive dis-
course particles, focusing on issues such as formant structure and voicing/phonation characteristics. A comparison is also made with previous analyses [E. Thom, MA thesis, UCL (2005); R. Eklund, J. Int. Phonetic Assoc. 38, 235–325 (2008)] as well as audio samples from other localities within the North Atlantic region.

5aSC28. The surface wave model of phonation threshold pressure and physical models of the vocal fold mucosa: Theory and experiment. Lewis P. Fulcher and Ronald C. Scherer (Bowling Green State Univ., Bowling Green, OH 43403)

One of the simpler manifestations of the role of the vertical phase difference in transferring energy from the glottal airflow to the motion of the vocal folds is the surface wave model (SWM) developed by Titze in 1988. He predicted that the phonation threshold pressure (PTP) should decrease as the glottal half width decreases and as the vocal fold thickness increases, and increase as the tissue damping increases, if no vocal tract is present. Further, his treatment of the SWM predicts that for a given glottal half width, converging glottal shapes should have a higher PTP than diverging glottal shapes. Although some of the experiments done with physical models of the vocal fold mucosa support these predictions, the measurements of the angle dependence of the PTP [Chan et al., “Further studies of phonation threshold pressure in a physical model of the vocal fold mucosa,” J. Acoust. Soc. Am. 101, 3722–3727 (1997)] do not follow the expected trend. A recent re-examination of the SWM found that the diverging-converging question is sensitive to the value of the entrance loss coefficient. Values of the coefficient near 1.37 seem to remove the discrepancy with the data. [Work supported by NIH R01DC03577.]


This study examines the phonation contrast of Yi, a Tibeto-Burman language of Southwestern China with phonation as a phonemic dimension (traditionally described as tense versus lax contrast). The language has seven monophthongs as well as three tones (high, mid, and low). Tense versus lax phonation contrast can apply to all the vowels and two tones (mid and low) in the language. In this study, both acoustic and electroglottographic data were collected with minimal pairs for all combinations of tones and vowels. Extensive acoustic measures were done by VOICESAUCE, including F0, H1, H2, H1-H2, H2-H4, H1-A1, H1-A2, F1, F2, and cepstral peak prominence (CPP); EGG measures were closed quotient (CQ) and peak-velocity (PV). Preliminary data suggest that H1-H2, CPP, CQ, and PV measures can exclusively distinguish the two phonation types. Moreover, EGG data also suggest voice quality interacts with tones, but the patterns vary among speakers. This might due to the different strategies used by speakers to produce the tense versus lax contrast. In addition, formant frequency measures also show salient differences between tense and lax vowels; tense vowels generally have higher F1 than the lax counterparts. All of the facts suggest that phonation contrasts in Yi are realized across different phonetic dimensions.

5aSC30. Acoustic analysis of phonation and tone interactions in Mazatec. Marc Garelek (Dept. of Linguist., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095-1543, margarelek@ucla.edu)

Jalapa de Diaz Mazatec is unusual in contrasting three phonation types (breathy, modal, and laryngealized) and three tone heights: low, mid, and high, including tone contours. Using the acoustic measures H1-H2, H1-A2, and cepstral peak prominence (CPP), previous research has shown that the three phonation types on mid tones are best distinguished at the beginning of the vowel, with each phonation gradually tending toward breathiness [Blankenship (2002)]. The present study extends previous findings by investigating tone and phonation interactions in over 650 Mazatec words and for a large range of acoustic measures. The speech samples, from the UCLA Phonetics Archive, were produced by six female and eight male speakers and vary across the three phonation types and two tonal groups (non-low and low). Using the voice analysis program VOICESAUCE [Shue et al. (2009)], acoustic measures such as harmonic amplitudes, F0, energy, CPP, formants, and bandwidths were extracted. The results generally support previous findings. Several measures also show effects of tone on phonation. In addition, how the phonation types and tone groups differ in timing and the effect of preceding consonants on phonation were investigated. Results show that preceding aspiration changes the timing and degree of phonation. [Work supported by NSF.]

5aSC31. A spectral-slope compensated scale for measuring perception of vocal aperiodicity. Bruce R. Gerratt and Jody Kreiman (Div. of Head/Neck Surgery, UCLA School of Medicine, 31-24 Rehab Ctr., Los Angeles, CA 90095-1794)

The influence of sound intensity on listeners’ perception of pitch and of sound frequency on the perception of loudness are well known. Similar interaction effects occur in voice quality perception, but are not well understood. Previous studies suggest that listeners’ sensitivity to noise levels in voice depends on the shape of the harmonic spectrum: As the amount of high-frequency harmonic energy present in the voice spectrum increases, listeners require a corresponding increase in noise energy to perceive a constant noise level. These results indicate that perceptually meaningful measurement of spectral noise in voice requires derivation of a spectral-slope-compensated noise scale (similar to equal loudness curves). Based on 12 natural voices, series of stimuli will be created by modifying (1) the noise-to-harmonics ratio (NHR) for different source spectral slopes and (2) the source spectral slope for different NHR values. Just noticeable differences for each parameter will be plotted as a function of the other, to create a set of equal noise contours reflecting the perceptual interactions of these two acoustic attributes. [Work supported by NIH.]

5aSC32. Voice quality as a cue to location in a speaker’s fundamental frequency f0 range: A perceptual study of English and Mandarin. Jason B. Bishop (Dept. of Linguist., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095, j.bishop@ucla.edu)

Recent research shows that listeners can locate a specific f0 within a speaker’s individual range without prior experience with the speaker’s voice [Honorof and Whalen, J. Acoust. Soc. Am. 117, 2193–2200 (2005)], suggesting that perhaps some acoustic parameters co-vary with f0 range and are available to listeners. Here we tested Honorof and Whalen’s hypothesis that this information lies in voice quality. English- and Mandarin-speaking listeners heard brief, steady-state /a/ tokens from various locations throughout the f0 ranges of ten English and ten Mandarin voices and were asked to identify where in the speaker’s pitch range a given token came from. To test for listeners’ use of voice quality in the task, we compared mixed effect linear models of responses that included three acoustic measures of the stimuli associated with voice quality: H1-H2, H2-H4, and cepstral peak prominence. Preliminary findings show (1) each of the voice quality measures contributed significantly to the models, although they accounted for only a small portion of the variance in responses compared to f0 and (2) models improved significantly when the language of the voice, but not of the listener, was included; this variable represents further language-specific properties of the voices still to be explored.

5aSC33. Online monitoring of auditory feedback is sensitive to language experience. Zhaocong Chen, Peng Liu (Dept. Rehabilitation Medicine, Sun Yat-Sen Univ., Guangzhou 510080, China), Emily Q. Wang (Rush Univ. Medical Ctr., Chicago, IL 60612), Charles R. Larson (Northwestern Univ., Evanston, IL 60208), and Hanjun Liu (Sun Yat-Sen Univ., Guangzhou 510080, China)

It has been demonstrated that vocal responses to pitch perturbations vary as a function of stimulus parameter and can be modulated according to the specific demands of vocal task. The purpose of this cross-language study was to examine whether the online monitoring of auditory feedback is sensitive to language experience during vowel phonation. Native speakers of Cantonese and Mandarin participated in the experiments. They were asked to vocalize a vowel sound /a/ at their conversational pitch, during which their voice pitch feedback was unexpectedly shifted (±50, ±100, ±200, or ±500 cents, 200 ms duration) and fed back to them over headphones. The results showed that, as compared to previous findings in English speakers [Chen et al. (2007)], both Mandarin and Cantonese speakers produced smaller but faster responses to pitch perturbations. In addition, Mandarin speakers produced larger response magnitudes than Cantonese speakers, and the modulation of response magnitudes as a function of stimulus magnitude was observed in Cantonese but not in Mandarin speakers. These findings demonstrate that voice F0 control is language dependent. Further, the diff-
different patterns of vocal responses between Mandarin and Cantonese speakers indicate that this highly automatic feedback mechanism works within the specific tonal system of each language.

**5aSC34. The role of creaky voice quality in Cantonese tonal perception.** Hiu-Wai Lam and Kristine M. Yu (Dept. of Linguist., Univ. of California, Los Angeles, 3125 Campbell Hall, Los Angeles, CA 90095, hiuwait208@ucla.edu)

Vance (1976) found a response bias against Tone 4 (mid-low falling) in a tonal perception experiment in Cantonese where synthesized stimuli varied only in F0, and Vance (1977) explained this by suggesting that creaky voice quality is a redundant cue for Tone 4. Indeed, there is evidence that creaky voice quality plays a role in tonal perception: in Mandarin, a language where creak is well-known to be a redundant cue for one of the tones, Tone 3 (low fall-rise), Belotet-Grenié and Grenié (1994) found that creaky instances of Tone 3 were recognized faster than non-creaky instances. The effect of creaky voice quality on the perception of Tones 4 and 6 (mid-low level) in Cantonese will be investigated using a 2AFC identification task of natural stimuli that were elicited in isolation and in connected speech. Variability in creak in the realization of Tone 4 occurred naturally in the elicited stimuli. If creaky voice quality plays a role in tonal perception, we hypothesize that overtly audible creak, as well as low H1-H2, a spectral index of creak, will bias listeners toward identification as Tone 4, as measured by d’ scores, and may also speed Tone 4 recognition.

**5aSC35. Perceptual similarities between native and non-native tones.** Xianghua Wu (Dept. of Linguist., Simon Fraser Univ., 8888 Univ. Dr. Burnaby, BC V5A 1S6, Canada, xianghua_wu@sfu.ca)

This study investigated the effects of L1 and L2 experience on the perceptual assimilation of non-native tones by native Mandarin or Thai listeners. Of these, 32 had 0.5–1.5 years of L2 learning experience while 40 had none. All listeners participated in a tonal assimilation task in which they first identified which tone in Mandarin or Thai sounded most similar to the Thai or Mandarin tone they heard, and then rated its goodness on a five point Likert scale. Stimuli included four Mandarin tones (high level, rising, falling-rising, and rising) and five Thai tones (mid, low-, falling-, high-, and rising tones) on four monosyllables: ʔuatu, ʔeiʔj, khua, ʔsaai/, and a hum. The Thai listeners, regardless of L2 experience, assimilated more L1 tone categories to L2 than did the Mandarin listeners. Nonetheless, the experienced listeners from the two languages showed a high degree of consistency in terms of which tones were assimilated. The perceived similarities between native and non-native tones were not always predictable from their acoustic similarities, and varied with L1 and L2 experience. Results will be discussed in terms of some well-known perceptual similarity measures, such as PAM. [Work supported by GIS.]

**5aSC36. Production and perception of lexical tones in Beijing and Taiwan Mandarin.** Ching-Yun Chang and Robert Allen Fox (SPA Labs, Speech and Hearing Sci. Ohio State, 1070 Carmack Rd., Columbus OH 43210-1002, chang.553@osu.edu)

The acoustic properties of four lexical tones between two regional varieties of Mandarin Chinese, Beijing Mandarin (Putonghua), and Taiwan Mandarin (Guoyu) were examined in terms of duration, pitch contours, and rms amplitude. Tokens included CV and V monosyllables, representing each of the four tones of Mandarin Chinese, and were produced in isolation and in sentence context by 15 adult native speakers. A different durational pattern of citation tones emerged for two dialect variety groups. In Taiwan Mandarin, it was T2>T1>T3>T4; whereas, it was T3>T2>T1>T4 in Beijing Mandarin [Deng et al. (2006)]. The durational discrepancy in isolated T3 may be related to different realization of T3 between two dialect groups. While T3 exhibited a falling-rising pitch contour in Beijing Mandarin, it was falling in Taiwan Mandarin. Such dialectal divergence in T3 contour shapes can be verified from the amplitude contours. Furthermore, pitch contours of other tones in two dialects will be compared to see if there is tonetic sound change in other Taiwan Mandarin tones. These surface acoustic variations in linguistically identical categories can result in perceptually ambiguous tones. A gating experiment was utilized to examine how native and non-native listeners adapt to speaker and dialect variability in the stimuli.

**5aSC37. Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects.** Hyunjung Lee and Allard Jongman (Dept. of Linguist., Univ. of Kansas, 1541 Lilac Ln., Lawrence, KS 66044, elvisdj@ku.edu)

The three-way laryngeal distinction among voiceless Korean stops has been well documented for the Seoul dialect. The present study compares the acoustic and aerodynamic properties of this stop series between two dialects, non-tonal Seoul and tonal South Kyungsang Korean. Sixteen male Korean speakers (eight from Seoul and eight from Kyungsang) participated. Measures collected included VOT, f0 at vowel onset and midpoint, H1-H2, and air pressure and airflow. The presence versus absence of tone affects both the acoustic and aerodynamic properties. First, Seoul speakers primarily use f0 to distinguish the three laryngeal gestures of Korean stops, while Kyungsang speakers mainly use VOT. Second, the high versus low tonal contrast for Kyungsang speakers makes f0 an unreliable acoustic cue for the three Korean stops. Third, dialectal differences in VOT to mark the three-way distinction support the notion of a diachronic transition whereby VOT differences between the lenis and aspirated stops in Seoul Korean have been decreasing over the past 50 years. Finally, the aerodynamic results make it possible to postulate the articulatory state of the glottis. Based on the acoustic and aerodynamic results, different phonological representations for the tonal and non-tonal dialects are suggested.

**5aSC38. Temporal changes in Mandarin tone due to speaking rate variation.** Joan A. Sereno and Hyunjung Lee (Dept. Linguist., Univ. of Kansas, 1541 Lilac Ln., Lawrence, KS 66045, sereno@ku.edu)

The present research investigates the effects of variation in speaking rate on the production of Mandarin tone. Fourteen speakers (6M, 8F) produced 15 syllables, each with the four different Mandarin tones. Speakers produced these syllables at fast, normal, and slow speaking rates in isolation. To induce rate change, the 60 target words were presented on a computer screen at different tempos. Consonant, vowel, and syllable duration were measured as well as turning point (TP) for the contour tones, Tones 2 and 3. TP is the time interval from tone onset to the lowest f0 value. Duration data of each tone revealed significant ambiguity across tones due to speaking rate. Specifically, for the contour tones, results indicated significant overlap in terms of duration and, as speaking rate increased, the TP for Tone 2 decreased to a lesser extent as compared to Tone 3. Moreover, the TP values for Tones 2 and 3 showed a sizable range, with the ambiguous region of overlap observed not at the faster but at the slower speaking rates. These findings will be discussed in terms of the temporal adjustments that occur in production when speaking rate changes. [Research supported by NSF.]

**5aSC39. An acoustic and electroglossographic study of Cantonese tone.** Kristine M. Yu, Hiu-Wai Lam, and Shing-Yin Li (Dept. of Linguist., Univ. of California, Los Angeles, 3125 Campbell Hall, Los Angeles, CA 90095, krisyu@humnet.ucla.edu)

Cantonese is a tone language with six tones (traditionally described as high level, high rising, mid level, mid-low falling, mid-low rising, and mid-low level), and also three “stopped” tones in CVC words, not studied here. The interest of this study was (1) how different acoustic cues can be used to classify the different tones, particularly in connected speech and (2) the interaction of voice quality and tone in Cantonese. Twelve speakers were recorded producing minimal pairs as in Wong ([2006]) in isolation, in isolated disyllables, and in sentence-medial disyllables, where the disyllables ranged over all possible 36 bitones in the language; all speakers also made electroglossographic (EGG) recordings. Acoustic measures were F0, F0′, cepstral peak prominence, and harmonic amplitudes H1 and H2, H1′-H2′, H1′-A1′, H1′-A2′, H1′-A3′, and H2′-H4′. EGG measures were closed quotient and peak of increasing contact. Measures were made automatically using VOICESAUCE and EGGWORKS. Preliminary results indicate variability in
the presence of creak in the realization of the lowest tone (mid-low falling) as well as variability in the presence of creak at the low F0 regions of the rises.

5aSC40. Identification of prominence and discrimination of pitch patterns by Japanese and American listeners. Irina A. Shport and Susan Guion-Anderson (Dept. of Linguist., Univ. of Oregon, 1290 Eugene, OR 97403, ishport@uoregon.edu)

This study examines how native language shapes the perception of prominence in three-syllable nonce words nenema with F0 patterns varying in F0 peak alignment and F0 fall. F0 is the fundamental cue to perception of pitch accent in Japanese, in which the accent location and accent type (accented versus unaccented phrases) are mainly defined by the relationship between the F0 peak, which is associated with the accented syllable, and the F0 fall, which follows the peak. In English, F0 fall is not considered to be a cue to stress. In two experiments, the alignment of the F0 peak (eight locations) and the F0 fall (no fall, moderate fall, steep fall) were manipulated. In the identification task, participants were asked to decide whether the first or the second syllable sounded more prominent to them. In the discrimination task, participants decided whether the pitch patterns of two words were the same or different. Japanese listeners were expected to be more sensitive to the F0 fall than American listeners for all F0 peak locations. Difference in the perception of peak location between the groups was also predicted, as a reflection of peak delay typical to the native language.

FRIDAY MORNING, 23 APRIL 2010

ESSEX A/B/C, 7:35 A.M. TO 12:00 NOON

Session 5aSP


R. Lee Culver, Cochair

Pennsylvania State Univ., Applied Research Lab., P.O. Box 30, State College, PA 16804-0030

Brett Bissinger, Cochair

Pennsylvania State Univ., Applied Research Lab., P.O. Box 30, State College, PA 16804-0030

Chair’s Introduction—7:35

Invited Papers

7:40

5aSP1. Statistical pattern classification: A review and some current problems/paradigms. David J. Miller (Dept. of Elec. Eng., The Penn State Univ., Rm. 227-C EE West Bldg., Univ. Park, PA 16802)

Unsupervised clustering is the ability to automatically partition a set of data patterns into meaningful groups without prior knowledge of the groups or their number. Supervised classification is the ability to automatically recognize a data pattern as an instance from one of a known set of classes. These problems are fundamental to a variety of application domains, including scientific (e.g., bioinformatics), engineering (e.g., speech recognition), business (e.g., marketing and document clustering), and military (target detection). In this talk, we first review the basic supervised and unsupervised classification problems, standard solution methodologies, and how to characterize their performance. We then discuss some recent problem variants and associated paradigms, including semi-supervised learning, unsupervised clustering in high-dimensional feature spaces, and decision fusion techniques.

8:00

5aSP2. Target/clutter discrimination using Bayesian active/passive data fusion. Brian R. La Cour, Jason M. Aughenbaugh, Bryan A. Yocom, and Thomas W. Yadichak (Appl. Res. Labs., The Univ. of Texas at Austin, P.O. Box 8029, Austin, TX 78713-8029)

Discriminating targets of interest from background clutter is a key challenge in underwater active sonar. This is particularly true in littoral areas, where scattering from the bottom, motion in the water column, and activity on the surface all contribute to produce echoes which are difficult to distinguish from targets of interest. Statistically, these effects manifest themselves in non-Rayleigh, heavy-tailed amplitude distributions. An approach to target/clutter discrimination is described which uses the complementary information from both active and passive acoustic sensors to facilitate this task. The method uses an efficient, grid-based Bayesian track-before-detect scheme to combine data from the two types of sensors by carefully modeling the effects of array processing, replica correlation, normalization, and clutter statistics. Representing measurements and uncertainty in terms of likelihood functions then provides a common framework for fusion. In this manner, active returns with coincident and appropriate passive signals are given more credence, while the presence of only one or the other, while perhaps suggestive, is not as compelling. The presentation will give a general overview of the approach, and an example will be used to illustrate the potential power of using passive data to mitigate active clutter. [Work supported by the Office of Naval Research.]
5aSP3. Moments in history. Leon Cohen (Dept. of Phys., Hunter College of CUNY, 695 Park Ave., New York, NY 10065)

We discuss how moments and functions of moments have been historically used to characterize and classify distributions, as well as to construct distributions. Among the methods that have been devised are the Gram-Charlier and Edgeworth series, which are methods used to improve on the Gaussian distribution when the moments deviate from Gaussian. We show how these series can be generalized and how any two distributions can be transformed into each other by systematically changing the moments. In this manner, one can develop families of distributions. We illustrate the methods by applying them to pulse propagation in a dispersive medium where the moments can be deterministic or stochastic. [Work supported by the Office of Naval Research.]

5aSP4. Computationally enabled alternatives to Gaussian classification and tracking—A survey. John R. Sacha (Appl. Res. Lab., Penn State, P.O. Box 30, State College, PA 16804, jsr9@psu.edu)

The normal distribution is a good model of physical phenomena in numerous situations where central limit theorem effects come into play. It has a prominent place historically because the many special properties of the Gaussian PDF often yield mathematically elegant closed-form solutions to problems. This role was especially important when numeric calculation was expensive. Normality is not a valid assumption in many applications of modern acoustical signal processing, perhaps especially involving classification and tracking. For pattern recognition, Gaussian classifiers are a straightforward methodology requiring mean and covariance estimates of the class distributions. However, in many cases, feature PDFs themselves are not multivariate-Gaussian; even if they are, estimating parameters in large feature spaces with limited data is difficult. In tracking, the Kalman filter is the canonical formulation, but it requires normally distributed process noise and measurement errors. The advent of cheap computation power has made practical a host of alternative models. In classification, this includes empirical PDF estimation via binning, fuzzy logic, neural nets, support vector machines, featureless classification, Bayesian belief nets, and decision trees. For tracking, particle filters are used to model densities as clouds of point estimates; although computationally expensive, they admit arbitrary densities and are amenable to parallel implementation. [Portions of this work were sponsored by ONR.]


A fundamental problem that arises in many applications is to estimate an unknown probability distribution from a set of observations drawn according to this distribution. Of particular interest is the case of non-parametric estimation in which little or nothing is known about the underlying distribution. Surprisingly, consistent estimators can be obtained even in this case and the resulting methods are called universal. A related problem is to estimate various information measures based on observations drawn one or two unknown distributions. This talk provides an overview of universal density estimation and the problem of estimating information measures. Some recent results and applications are also described.

5aSP6. Classification of non-Gaussian reverberation as a mixture. Nicholas P. Chotiros (Appl. Res. Labs., The Univ. of Texas at Austin, 10000 Burnet Rd., Austin, TX 78758, chotiros@arlut.utexas.edu)

There are a few different ways to obtain non-Gaussian statistics, one of which is through the heterogeneity of the seafloor. It is often the case that the seafloor is a patchwork of different bottom types. Measurements of seafloor reverberation, using omni-directional sources and receivers, often show Gaussian statistics, because it is the result of the superposition of a large number of random contributions, satisfying the central limit theorem. For sonar systems that have the spatial resolution to resolve the patches, the statistics of the reverberation will change from one resolution cell to the next. Due to positional inaccuracies associated with most sonar systems, it is often not possible or feasible to separate the reverberation from each bottom type. When taken as a whole, the reverberation will have non-Gaussian statistics. The statistical properties will be explored. [Work supported by the Office of Naval Research, Ocean Acoustics Program.]
5aSP8. Active sonar clutter classification using higher order moments. James M. Gelb and Andrew W. Oldag (Appl. Res. Labs., Univ. of Texas at Austin, 10000 Burnet Rd., Austin, TX 78758, gelb@arlut.utexas.edu)

The statistics of normalized matched-filter echoes from an active sonar system operating in a myriad of oceanic environments has been studied extensively for three broad clutter classes including using low-order cumulants to classify subregions of the data [Gelb et al., Proceedings of the ISURC (2008) and references therein]. That work compared empirical distributions to parametric models (e.g., the K distribution and the generalized Pareto distribution). A report on extensions of this work is presented including studies of the accuracy of analytic parameter estimation methods and the efficacy of using higher order moments in the classification process. For each class, with increasingly heavy non-Rayleigh distributed tails, comparisons are made of brute force parameter estimation with the use of analytic estimators. Additionally, comparisons of higher order moments (including skew and kurtosis) computed from the data are made with analytic fits to the data. Using a feature-based classifier, the gains of using increasingly higher order moments are assessed. [Work sponsored by the Office of Naval Research (ONR).]

10:20

5aSP9. Improving anti-submarine warfare tracking performance by incorporating classification information. William H. Mortensen, David W. Krout, and Jack McLaughlin (Appl. Phys. Lab, Univ. of Washington, 1013 NE 40th St., Box 355640, Seattle, WA 98105-6698, william55@u.washington.edu)

Prior research has shown that more accurate classification of contact amplitudes can improve tracking performance. In this work, we augment standard, contact-based trackers with a classifier run on features of the received time series from which the contacts were extracted. Ground truth information from benchmark datasets is used in a flexible simulation framework built around the sonar simulation toolset (SSST) to generate simulated target and clutter returns. The simulated returns are used as input to the classifier and the contacts from the benchmark datasets as input to the tracker. The classification information provides an additional input to the association step in the probabilistic data association (PDA) and joint PDA trackers, and to the probability of target for each contact in the Bayesian tracker. The results show that even relatively poor classification can make a noticeable improvement in tracking performance. [This work was funded by the U.S. Office of Naval Research, Contract No. N00014-01-G-0460, Delivery Order #36.]

11:00

5aSP10. Model-based detection of buried objects. Edmund J. Sullivan (EJS Consultants, 46 Lawton Brook Ln., Portsmouth, RI 02871, paddypriest@aol.com) and Ning Xiang (Rensselaer Polytechnic Inst., Troy, NY 12180)

An approach to the detection of buried objects is to excite the ground with low-frequency, high-energy sound waves, which then excites a resonance in the buried object. The ensuing vibration causes a detectable signal on the surface of the ground, which can be detected using a laser doppler vibrometer (LDV). The original detection technique used a sliding bandpass filter to process the scattered LDV energy, providing an energy map of the area scanned by the LDV, which indicates the location of the object. The performance of detection is often limited by speckle noise, a type of noise arising from the coherent nature of the laser beam. A more recent technique utilizes an autoregressive model of this noise. This leads to an inverse filter that whitens the noise. Upon the appearance of any target data in the signal, a whiteness test indicates a detection. This approach has demonstrated improvement over the bandpass filter approach. This paper demonstrates a further improvement by augmenting the prewhitener with a model of the mine itself. This provides significant improvement by both enhancing the mine signal and improving the detection performance. Experimental results are shown.

11:20

5aSP11. Sonar waveform design for detection of elastic objects. Brandon Hamschin (Dept. of ECE, Univ. of Pittsburgh, 348 Benedum Hall, Pittsburgh, PA 15261, bmh52@pitt.edu) and Patrick Loughlin (Univ. of Pittsburgh, Pittsburgh, PA 15261)

Animals that navigate and hunt by echolocation, such as some bats and marine mammals, have been observed to change their sonar pulse duration on the environment, as well as during hunting. It has become of interest to incorporate these strategies into man-made sonar waveform and receiver design. We examine the benefits of optimal waveform design versus transmitting a linear FM waveform for detecting elastic objects. Performance loss suffered by assuming a point target is also examined. Our approach utilizes a method recently proposed by Kay to design the optimal power spectrum of the transmit waveform. Because there is an unlimited number of waveforms with the same power spectrum, we further impose a time domain constraint, in terms of the signal duration, to obtain a unique optimal waveform. [Work supported by ONR 321US.]

11:40

5aSP12. Error bounds for classifying targets in non-Rayleigh clutter. Douglas A. Abraham (CausaSci LLC, P.O. Box 5892, Arlington, VA 22205, abraham@icee.org)

False alarms in active sonar systems are often represented statistically as having a probability density function (PDF) with tails heavier than the traditionally assumed Rayleigh PDF. Distributions such as the Weibull, K, and Poisson–Rayleigh have been used to represent such non-Rayleigh clutter and to derive the associated probabilities of false alarm and detection, the latter for standard target models (e.g., nonfluctuating and fluctuating). In this presentation, the probability of at least one misclassification per ping ($P_c$) is evaluated when classifying the standard target types against non-Rayleigh clutter having equal average power. The Bhattacharyya bound is used as an upper bound for the Bayesian error probability ($P_e$) of a classifier operating on a single cluster and combined with the average number of clusters per ping ($N_c$) to form $P_e$. As expected, $P_e$ increases as the clutter becomes more non-Rayleigh. Clutter statistics are seen to affect $P_e$ through both $P_e$ and $N_c$. However, the effect of clutter statistics on $N_c$ is significantly greater than on $P_e$. Increasing bandwidth reduces $P_e$ (and therefore $P_c$) when its impact is modeled as an increase in the number of independent samples available for classification. [Work sponsored by the Office of Naval Research under Contract No. N00014-09-C-0318.]