Session 5aAA

Architectural Acoustics: Topics in Architectural Acoustics

Benjamin C. Mueller, Chair
Ostergaard Acoustical Associates, 200 Executive Drive, West Orange, New Jersey 07052

Chair's Introduction—9:00

Contributed Papers

9:05

5aAA1. Investigation of early reflections on the perception of audio. John Gorr (Columbia College Chicago, 1025 W. Wood St., Palatine, IL 600067, jgorr1@msn.com)

The effects of early reflections in a nonsymmetrical control room were measured using time delay spectrometry and MLS methods. The tests revealed that comb-filtering patterns were the dominant source of spectral distortion. The mixing console reflections were investigated to reveal the subjective impact of the comb-filtering at the mix position. Tests were performed with and without acoustic foam on the console to determine the influence of the early reflections and then compared with theoretical values for the filtering patterns. Results were compared and used to qualify the subjective perception of the distortion.

9:20

5aAA2. Sound absorption and reflection from ceilings in open offices. Alf Warnock (Natl. Res. Council Canada, M59, 1200 Montreal Rd., Ottawa, ON K1A 0R6, Canada, alf.warnock@nrc-cnrc.gc.ca)

Sound attenuation between work areas in open offices depends on barrier height and the sound-absorbing properties of the ceiling. Two methods for evaluating the sound-absorption ceilings are available: ASTM E1111, which gives a rating called the articulation class (AC), and ASTM C423, which gives sound-absorption average (SAA). Measurements of six ceiling systems made in accordance with ASTM E1111 and ASTM C423 provide a link between AC and SAA. Manufacturers often do not provide both ratings for their products, and it is useful to be able to use either rating system. SAA and measured AC for a 1.5-m-high screen are related by $AC=102\times SAA+91.4$. SAA and calculated values of AC for a 1.8-m-high screen are related by $AC=118\ SAA+93$. In the absence of E1111 test data, C423 SAA values can be used to select ceiling systems. The commonly recommended minimum SAA of 0.9 corresponds to an AC of 180 when the screen height is 1.5 m and to an AC of 200 when the screen height is 1.8 m. An approximate conversion from absorption coefficients to reflection coefficients agrees well with earlier work.

9:35

5aAA3. Prediction of broadband nonuniform time-dependent acoustic fields in enclosures with diffuse reflection boundaries using energy-intensity modes. Donald B. Bliss and Linda P. Franzoni (Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708, dbb@duke.edu)

A new analysis of high-frequency broadband reverberant sound fields in rooms with diffuse reflection boundaries is described. Depending on shape, source location, and distribution of wall absorption, rooms exhibit spatial variation in steady-state mean-square pressure and also spatial dependence of decay time characteristics. The room boundaries can be replaced by a distribution of uncorrelated broadband directional energy-intensity sources. In steady state with diffuse reflection boundaries, the interior pressure field produced by these sources satisfies Laplace's equation. The mean-square pressure field is expressed as a sum of constituent modes. The intensity field, which is related to the pressure field in a complex way, can be calculated for each mode. Boundary conditions relate averaged intensity and pressure. The mean-square pressure is expressed in terms of the modal sum. Lower order modes are responsible for the overall smooth spatial variation in the reverberant field; higher modes account for more rapid local variations near walls due to changes in wall properties. In the transient problem, the spatial eigenmodes decay at different rates, leading to a spatial redistribution of reverberant energy, and causing the decay curves to be a function of position in the room. Sample calculations and comparisons with other solution methods are presented.

9:50

5aAA4. An acoustical analysis of a room with a concave dome ceiling element. Sentagi S. Utami (N283 Eyring Sci. Ctr., Brigham Young Univ., Provo, UT 84602, sentagi@yahoo.com)

Concave surfaces are often considered detrimental in room acoustics, especially because of the impact they have on the distribution of sound energy. This paper explores certain acoustical characteristics and anomalies found in spaces below concave dome ceiling elements. The architectural design of the Darusshollah mosque in East Java, Indonesia is used as a case study with specific spatial and functional concerns. Investigations of the mosque have been conducted through both a 1:12 scale model and a computer model that utilizes ray tracing and image source methods. Analysis techniques are discussed. Results are presented and compared to provide useful insights into the acoustics of such distinctive environments.

10:05

5aAA5. Investigation on the flanking transmission of impact sound insulation of floor. Giovanni Semprini and Alessandro Cocchi (DIEMCA Dept. of Eng., Univ. of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy)

Impact sound pressure level of floors depends, as general rule, on direct sound radiated by the floor excited by a standard tapping machine and on flanking transmission of lateral walls. Depending on the kind of junction between the floor and walls of the receiving room, flanking paths can be more or less important. Requirements of laboratory test specimens are not well specified in EN ISO 140 standards, particularly for junctions of the test floor and lateral walls. In this paper measurements performed at DIENCA laboratory are presented in order to evaluate the influence of flanking transmission on impact sound pressure levels of a standard floor and on impact sound reduction level of standard floor with a resilient layer. Measurements are performed on a 14-cm-thick concrete floor in two different conditions: first connected on two sides of the receiving room and then on all four sides. Impact sound pressure levels and vibration levels are analyzed in order to evaluate the contribution of different transmission paths. As the measurements were carried on in the new facility for measurement of the contribution of the flanking transmission, this facility will be exhaustively presented in the paper.

9:25

5aAA6. Acoustics of early music spaces from the 11th to 18th century: Rediscovery of the acoustical excellence of medium-sized rooms and new perspectives for modern concert hall design. Alban Bassuet (Arup Acoust., 155 Ave. of the Americas, New York, NY 10013, alban.bassuet@arup.com)

The acoustical characteristics of 50 rooms that played a prominent role in the history of music between the 11th and 18th centuries were studied. The rooms include basilicas, oratorios, organ churches, and the great halls and courts of the European palaces. The research provides an understanding of the acoustical features that suit the early music repertoire, and how these rooms achieved an enhanced emotional engagement through their unique acoustical characteristics. This paper provides a summary of the acoustic measurements, which include binaural and B-format recordings in each of the rooms, and presents a unique new approach to understanding their subjective characteristics through detailed analysis and auralization of their 3-D impulse response. The study shows that the timing and direction of reflections in three dimensions is critically important to defining the subjective characteristic of a room. The results emphasize the importance of developing techniques to understand the 3-D impulse response and using auralization techniques for interpreting results and making subjective judgments. The enhanced musical experience that is achieved in these early rooms offers an invitation to rethink modern acoustics and to develop a new design approach that focuses more strongly on the subjective response and emotional engagement of the music.

10:35

5aAA7. Use of surrogate samples to study variation of absorption coefficients of fiberglass with altitude. Richard D. Godfrey (Owens Corning, Sci. and Technol., 2790 Columbus Rd., Granville, OH 43023)

ASTM C 423 identifies air temperature and relative humidity as significant parameters, but does not address air density effects. At constant temperature, air density decreases approximately 20% from sea level to 5000 ft altitude. In previous papers, normal and diffuse field analysis showed significant changes in predicted absorption coefficients with altitude. These predictions were validated experimentally for normal incidence by making measurements in a vacuum chamber. Reverberation chambers cannot withstand depressurization. They also exhibit significant interlaboratory measurement variability. Another method was soot. The Mechel design charts are normalized by two parameters. One is not dependent on air density. The other is the ratio of flow resistance and the impedance of air. If thickness is held constant, the effect of lowering air density can be studied by increasing the sample flow resistivity of the sample. This surrogate sample should emulate absorptive performance at high elevations in sea level laboratories. Impedance tube measurements using surrogate samples emulated the effects observed in the pressure chamber study. The next step is to use surrogate samples to investigate air density effects in diffuse fields using the ASTM C 423 test method in a single laboratory.

10:50

5aAA8. Acoustical phenomenon in ancient Totonac's monument. José Sánchez-Dehesa, Andreas Håkansson (Nanophotonic Technol. Ctr. and Dept. of Electron. Eng., Polytechnic Univ. of Valencia, E-46022 Valencia, Spain), Francisco Cervera, Francisco Meseguer (Polytechnic Univ. of Valencia, E-46022 Valencia, Spain), Betsabé Manzanares-Martínez, and Felipe Ramos-Mendieta (Univ. of Sonora, Hermosillo, Sonora 83190, Mexico)

The circle of gladiators is a monument built by Totonac Indians in the ceremonial site of Cempoala, which is located near Veracruz (Mexico). The city is believed to date to around 1200 A.D. The monument is a round structure with crenellated wall tops, and it has a diameter of 13.4 m. Though the deterioration of this monument is noticeable, it presents a singular acoustical phenomenon whose strength had to be probably extraordinary on the date of its construction. In brief, along any diameter in the circle, one can find two focal points such that if one person speaks on one focus, another person located on the other hears the sound reinforced. In other words, this circular place acoustically behaves as if it were elliptical. Here, we report the experimental characterization of the phenomenon and present a theoretical explanation. Also, the intentionality of the Totonacs is speculated since these people are associated with the Mayan culture, which is known by its realizations of environments with astonishing sonic properties. [Work supported by CEAL-UAM of Spain.]

FRIDAY MORNING, 28 MAY 2004 NEW YORK BALLROOM A, 8:00 TO 10:45 A.M.

Session 5aAOa

Acoustical Oceanography and Animal Bioacoustics: D. Van Holliday Special Session on Acoustical Measurements of Marine Organisms III

John K. Horne, Chair

School of Aquatic and Fisheries Science, University of Washington, Box 355020, Seattle, Washington 98195-5020

Contributed Papers

8:00

5aAOa1. Twenty-five years with Van Holliday in the development of high-frequency technology and analysis algorithms to measure zooplankton distributions. Richard E. Pieper (Southern California Marine Inst., 820 S. Seaside Ave., Terminal Island, CA 90731, pieper@usc.edu)

Initial studies using high-frequency acoustics at four individual frequencies (0.5–3.0 MHz) were begun in the 1970s to measure acoustical scattering from zooplankton. Acoustical measurements were made at sea by profiling vertically in the water column. Zooplankton were collected, identified and measured, and target strength measurements were made on individual zooplankton in the laboratory. Concurrently, various acoustical scattering models were analyzed to enable the calculation of the size-frequency distribution of zooplankton from the acoustical data. A multi-frequency acoustic profiling system (MAPS) was then developed (21 different acoustical frequencies). This system was used to measure oceanic structure off of southern California, plumes off of central California, Gulf Stream features, and oceanic structure in the Irish Sea. Analyses of these data indicated that 21 frequencies were more than needed. Four to six frequencies were adequate for most studies, and the Tracor Acoustic Profiling System (TAPS) was developed. This system has been used and modified for a wide variety of studies. These studies range from large-scale patterns of zooplankton in the Arabian Sea to the measurement of thin layers in many different oceanic systems. The use of these systems now provides us with high-resolution measurements of zooplankton distributions in the sea.

8:15

5aAOa2. Acoustic backscatter models of fish: Gradual or punctuated evolution. John K. Horne (Univ. of Washington, School of Aquatic and Fishery Sci., Box 355020, Seattle, WA 98195, jhorne@u.washington.edu)

Sound-scattering characteristics of aquatic organisms are routinely investigated using theoretical and numerical models. Development of the inverse approach by van Holliday and colleagues in the 1970s catalyzed the development and validation of backscatter models for fish and zooplankton. As the understanding of biological scattering properties increased, so did the number and computational sophistication of backscatter models. The complexity of data used to represent modeled organisms has also evolved in parallel to model development. Simple geometric shapes representing body components or the whole organism have been replaced by anatomically accurate representations derived from imaging sensors such as computer-aided tomography (CAT) scans. In contrast, Medwin and Clay (1998) recommend that fish and zooplankton should be described by simple theories and models, without acoustically superfluous extensions. Since van Holliday's early work, how has data and computational complexity influenced accuracy and precision of model predictions? How has the understanding of aquatic organism scattering properties increased? Significant steps in the history of model development will be identified and changes in model results will be characterized and compared. [Work supported by ONR and the Alaska Fisheries Science Center.]

8:30

5aAOa3. A nonlinear model-based acoustic inversion to estimate the abundance and biomass distributions of marine organisms. Dezhang Chu and Peter Wiebe (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Multi-frequency and/or broadband acoustic systems can be used to estimate the abundance and biomass distributions of marine organisms with inversion techniques [Holliday *et al.*, J. Cons. Int. Explor. Mer. **46**, 52–67 (1989)]. The linear inversion scheme is commonly used in fisheries and zooplankton acoustics and is based mostly on relatively simple scattering models. A more advanced nonlinear inversion method is presented in this talk. It uses the more sophisticated scattering models and the material properties measured *in situ* to estimate the abundance and biomass of marine organisms. A multi-frequency acoustic data set collected with the bio-optical multi-frequency acoustical and physical environmental recorder (BIOMAPER-II) during the austral fall Southern Ocean GLOBEC broad-scale cruise in 2002 was used in the inversion. At the depth where the Antarctic krill (*Euphausia superba*) were the dominate scatterers, the behavioral information of the animal such as mean angle of orientation and the standard deviation of the tilt angle were estimated with the nonlinear inversion. Furthermore, nonuniqueness and uncertainty inherently associated with the nonlinear inversion are analyzed. [Work supported by the NSF.]

8:45

5aAOa4. High-frequency acoustic volume scattering from zooplankton and moving oceanic microstructure. Andone C. Lavery, Peter H. Wiebe, Raymond W. Schmitt, Timothy K. Stanton, Tetjana Ross, Gareth Lawson, Nancy Copley (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Karen E. Fisher (Los Alamos Natl. Lab., Los Alamos, NM 87545), and Fabian Wolk (Rockland Oceanogr. Services, Inc., Victoria, BC V9A 4B6, Canada)

It is well accepted that high-frequency acoustic scattering techniques can be used to perform rapid, synoptic surveys of fish and zooplankton over relevant spatial and temporal scales. However, the use of these remote sensing techniques to probe small scale physical processes, such as oceanic microstructure, has not been fully accepted, or exploited. Yet there is a growing body of evidence in the form of both laboratory and field measurements suggesting that their use is feasible. Discrimination of scattering from microstructure versus zooplankton, which span similar spatial scales, is typically a limiting factor. Currently, acoustic discrimination of turbulence from zooplankton relies on either source of scattering being dominant as well as on the availability and accuracy of scattering models. A model for scattering from oceanic microstructure that includes fluctuations in the density and sound speed is presented. The effects on scattering from a layer of microstructure with a mean fluid velocity are discussed. Backscattering predictions are made based on data collected in Hudson Canyon with a tethered free-falling high-resolution vertical microstructure profiler, and compared to two-frequency acoustic data (120 and 420 kHz). The contribution to scattering from zooplankton is also estimated from nearby depth-resolved net tows together with zooplankton scattering models.

9:00

5aAOa5. Ocean acoustic backscattering: When you can ignore acoustic scatter from turbulence and when you can't. Tetjana Ross, Andone Lavery (Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Rolf Lueck (Univ. of Victoria, Victoria, BC V8W 2Y2, Canada), Peter Wiebe, and Gareth Lawson (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

While models predicting measurable levels of acoustic backscattering from oceanic turbulence have been around for decades, they have proven notoriously hard to confirm. This is, in part, because potential turbulent scattering layers often coincide with zooplankton layers. Therefore, as zooplankton are known to cause measurable acoustic scatter, the source of scatter is obscured. Furthermore, estimates of zooplankton abundance from acoustic scattering measurements have been shown to agree with independent measures under a number of circumstances. This gives circumstantial evidence to fuel the belief that scatter from turbulence is negligible. In addition, even if the turbulent scattering theory is correct, it predicts that over most of the ocean the turbulent intensities and/or stratifications are too weak to give turbulent scatter of a similar magnitude as is observed from zooplankton layers. Yet, despite all this, here data are presented that show circumstances when scatter from turbulence is significant. The areas of the ocean (mostly coastal) that one might have to be concerned about turbulent scatter are discussed, as well as the intriguing idea of using technology already developed for zooplankton observation to measure turbulence. This technology could facilitate the *in situ* study of biophysical interactions between zooplankton and turbulence.

9:15

5aAOa6. Effect of orientation of euphausiids and copepods on acoustic target strength: Implications for measurements from down-looking and side-looking acoustic systems. Malinda Sutor and Timothy J. Cowles (College of Oceanic and Atmospheric Sci., Oregon State Univ., 104 Ocean Admin. Bldg., Corvallis, OR 97331, msutor@coas.oregonstate.edu)

Multifrequency acoustics is a potentially useful tool for zooplankton ecologists to rapidly map distributional patterns and determine taxonomic and size composition of scatterers. This is typically done with down-looking or side-looking acoustic systems. The ability to estimate taxonomic and size composition acoustically depends upon accurate models of scattering from zooplankton. In this study, distorted wave Born approximation (DWBA) models of individual zooplankton were used to illustrate that the frequency dependence of expected scattering is different for down-looking and side-looking acoustic systems due to the differences in orientation of scatters within the sampled volume. The results show that peaks and nulls occur at different frequencies for each system with maximum target strength differences of 30 dB. The models were used to predict volume scattering (S_V) based on zooplankton collected from MOCNESS tows. Predicted S_V was compared to S_V measured by both a down-looking HTI acoustic system and a side-looking TAPS. The results showed that changes in orientation can have large effects on predicted total S_V, particularly at higher frequencies, and demonstrate that choice of orientation parameters is important and different parameters should be used when comparing predicted S_V with measured S_V from down-looking or side-looking acoustic systems.

9:30

5aAOa7. Comparing high-frequency scattering by a fish swimbladder and a gas-filled ellipsoid. Kenneth G. Foote (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543) and David T. I. Francis (Univ. of Birmingham, Birmingham B15 2TT, UK)

High-frequency backscattering spectra have been computed for two air-filled sacs at 51-atm ambient pressure by means of the boundary-element method. One applies to the swimbladder of a 39-cm-long specimen of pollack (*Pollachius pollachius*), with known morphometry expressed in a 3181-node mesh. The other applies to a prolate ellipsoid, whose major axis is that of the mapped swimbladder, 14.1 cm, and whose minor axis, 1.24 cm, has been determined so that the volume of the ellipsoid is essentially identical to that of the mapped swimbladder. The dorsal-aspect backscattering cross section of each sac has been computed at each frequency over a normal distribution of tilt angles, measured relative to the longitudinal or major axis, with mean −4.4 deg and standard deviation 16 deg. Computations have been performed over the frequency range 20–40 kHz in increments of 25 Hz. The spectra are characterized and compared. [Work supported by ONR.]

9:45

5aAOa8. Some anomalous time domain back scattering phenomena from low contrast fluid spheres and cylinders. C. Feuillade (Naval Res. Lab., Stennis Space Ctr., MS 39529-5004), D. Chu (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and C. S. Clay (Geophys. and Polar Res. Ctr., Univ. of Wisconsin—Madison, Madison, WI 53706)

Anderson's scattering theory [J. Acoust. Soc. Am. **22**, 426–431 (1950)], and the corresponding formalism for an infinite cylinder, have been used to calculate the scattered impulse responses for low contrast fluid spheres and cylinders, where the density and sound speed of the object are fractionally higher than the surrounding medium. The returns from the front and rear faces of the object are individually identifiable. In the case of a sphere, both of these features are positive spikes. In the cylindrical case, the second feature is a negative first derivative of a spike. In both instances, the waveform of the second feature is different from that expected from simple physical reasoning, i.e., the raypaths and reflection coefficients. The predictions of the calculations are investigated by using them to analyze experimental scattering data from decapod shrimp near broadside incidence. By convolving the calculated impulse response with the incident acoustic signal, comparisons with the scattering data are made. Both the sphere and cylindrical calculations appear to fit the data well. The significance of these results, and possible explanations for the anomalous scattering from the rear face, are discussed. [Work supported by ONR, University of Wisconsin Weeks Fund.]

10:00

5aAOa9. Modeling of surficial sediment alteration by biology. Dajun Tang (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

This paper describes fine-scale measurements and modeling of surficial sediment roughness at a site in the East China Sea using a conductivity probe. The spatial resolution of the measurements is designed to obtain environmental data suitable for modeling acoustic backscatter in the mid-frequency (3–4 kHz) range. The power spectrum of the bottom roughness is estimated and it is found that bottom roughness is dominated by small features caused by bottom-dwelling organisms. This is confirmed by video images of the same spot of seafloor. A model is developed to simulate the random distributions of these bottom features. The model employs a superposition of discrete features, which result in a power spectrum that is consistent with the measured power spectrum. Potentially this kind of model can provide a remote sensing means to estimate bottom biological populations through measuring sound backscatter from the bottom. [Work supported by ONR Ocean Acoustics Code.]

10:15

5aAOa10. The Bergen multifrequency analyzer (BMA): A new toolbox for acoustic categorization and species identification. Egil Ona, Rolf Korneliussen, Hans Petter Knudsen (Inst. of Marine Res., P.O. Box 1870, 5817 Bergen, Norway, egil.ona@imr.no), Kjell Rang, Inge Eliassen, Yngve Heggelund, and Daniel Patel (Christian Michelsen Res. AS, 5892 Bergen, Norway)

Multifrequency split-beam echo sounders with nearly identical and overlapping acoustic beams have been regularly used in acoustic surveys for fish stock abundance estimation. Calibrated raw data from up to six simultaneously working echo sounders at 18, 38, 70, 120, 200, and 364 kHz were applied for developing a new processing tool for real-time acoustic target categorization and acoustic species identification. The system now handles raw data from the Simrad EK500 and EK60 split-beam echo sounders, and performs a stepwise, modular sequence of analysis, like bottom detection, noise quantification and removal, target categorization, and school detection in near-real time. Direct generation of new, synthetic echograms, based upon the measured frequency response of the targets, is also one of the most useful features of the system. This information may significantly increase the accuracy of acoustic survey estimates of fish and zooplankton. New routines for noise removal, target categorization, and school detection will be presented, as well as new methods for training and building the artificial experience of the analyzer.

10:30

5aAOa11. An echo analysis technique for estimating the fish population. C. P. Anil Kumar, Sajith N. Pai, N. Soniraj, M. H. Supriya, James Kurian, C. Madhavan, and P. R. Saseendran Pillai (Dept. of Electron., Cochin Univ. of Sci. and Technolgy, Cochin-22, Kerala, India)

The development of an algorithm for the estimation of biomass by acoustic remote sensing is presented in this paper. The distinctive features of the algorithm include the implementation of time-varied gain function, proper accounting of beam factor effects, implementation of backscattering levels for selected species and processing for echo-count as well as echo-integration. In order to optimize the backscattering levels, numerical analysis for geometrical backscattering of selected marine species were carried out and validated with *in situ* measurements. The biomass information of the concerned marine specie is made available by subjecting the backscattered raw data to a series of processes. The performance validation of the algorithm under *in situ* conditions yielded encouraging population estimation results and is being fine-tuned with the field data. This algorithm will provide an efficient technique to parametrically compute the target strength, leading to the estimation of the stock of commercially important selected marine species under varied environmental conditions in different regions.

Session 5aAOb

Acoustical Oceanography: Acoustical Oceanography Prize Lecture

Peter F. Worcester, Chair

Scripps Institution of Oceanography, University of California—San Diego, 9500 Gilman Drive, La Jolla, California 92093-0225

Chair's Introduction—10:55

Invited Paper

11:00

5aAOb1. Ocean acoustic inversion for seabed geoacoustic properties. Stan E. Dosso (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, sdosso@uvic.ca)

Estimating geoacoustic properties of the seabed from ocean acoustic data provides a convenient *in situ* alternative to direct sampling (e.g., coring), with parameter sensitivities relevant to sonar applications. However, this requires solving a strongly nonlinear inverse problem, which is inherently nonunique. Hence, quantifying uncertainties for the recovered geoacoustic parameters is an important, but challenging, problem. This talk will describe a nonlinear Bayesian approach to geoacoustic inversion based on estimating properties of the posterior probability density (PPD), which combines information from observed data with prior information. An efficient Markov-chain method (Gibbs sampling) is applied to extract properties of the PPD, including optimal parameter estimates, marginal probability distributions, variances/covariances, and interparameter correlations. The inversion formulation is general, and will be illustrated with examples for a variety of approaches to geoacoustic inversion, including the inversion of acoustic field data, seabed reflectivity measurements, acoustic reverberation data, and ambient noise measurements.

FRIDAY MORNING, 28 MAY 2004 RIVERSIDE SUITE, 8:00 TO 10:30 A.M.

Session 5aBB

Biomedical Ultrasound/Bioresponse to Vibration: Tissue Characterization and Gene Transfection

Subha Maruvada, Chair

FDA Center for Devices and Radiological Health, Rockville, Maryland 20852

Contributed Papers

8:00

5aBB1. Acquisition of transcranial ultrasonic signals from adult subject's applications to the detection of brain injury. Joel Mobley, Tuan Vo-Dinh (Oak Ridge Natl. Lab., P.O. Box 2008, Oak Ridge, TN 37831-6101, vodinht@ornl.gov), Brian J. Daley (Univ. of Tennessee Medical Ctr., Knoxville, TN 37920), and Martin C. Holland (San Francisco General Hospital, San Francisco, CA 94143)

In the clinical setting, several technologies (e.g., x-ray CT, MRI, color-duplex ultrasonography) can provide doctors with noninvasive windows through the cranium into the brain. In the field, however, there is a need for a portable device that can noninvasively detect brain injuries at the scene of a trauma, providing first responders with urgent diagnostic information in the critical first hour postincident. Such an instrument should be simple to operate and must supply results that are straightforward to interpret. This work is part of an effort to devise ultrasonic methods that can serve as the operating principle for such a device. In this talk, human studies are described in which backscattered ultrasound data were obtained transcranially from adult volunteers. The ultrasound signals, acquired with single transducers in pulse–echo mode, are interpreted anatomically with the aid of a multilayer model of the propagation path. Through the model, the complex structures of the echo patterns are revealed and the important roles of the scalp and skull layers are made evident. The ultrasound data are further analyzed to assess the suitability of these signals for implementing a tissue characterization approach to the injury detection problem.

8:15

5aBB2. Importance of *ka*-range on the simultaneous estimation of scatterer size and total attenuation from ultrasound backscattered waveforms. Timothy A. Bigelow and William D. O'Brien, Jr. (Dept. of Elec. and Computer Eng., Bioacoustics Res. Lab., Univ. of Illinois, Urbana, IL 61801, bigelow@uiuc.edu)

Considerable effort has been directed to diagnosing the malignancy of solid tumors noninvasively by determining a scatterer size from a statistical analysis of the ultrasound-backscattered waveforms. *In vivo* tumor assessments have had limited success due to frequency-dependent attenuation along the propagation path masking the frequency dependence of the scatterer size. In this study, both attenuation and size were solved simultaneously by a two-parameter minimization of the mean squared error between a reference spectrum, modified by the attenuation and scatterer size, and the backscattered waveforms. The performance of the approach was assessed by simulations of a homogeneous region containing scatterers with a Gaussian impedance distribution. The simulations varied the effective radius of the scatterers (5 to 150 mm), the attenuation of the

5a FRI. AM

region (0 to 1 dB/cm/MHz), and the bandwidth of the source. In all cases, comparable accuracy and precision of the scatterer size were obtained whenever the range of *ka* values (wavenumber times effective radius) had the same width (largest *ka* minus smallest *ka*). The precision and accuracy improved with increasing width. A width of 1 gave an accuracy/precision of $\sim 15\% \pm 35\%$ whereas a width of 1.5 gave an accuracy/precision of $\sim 5\% \pm 15\%$. [Work supported by a Beckman Institute Fellowship.]

8:30

5aBB3. Characterization and differentiation of three solid tumors using quantitative ultrasound. Michael L. Oelze, William D. O'Brien, Jr. (Dept. of Elec. and Computer Eng., Univ. of Illinois, 405 N. Mathews, Urbana, IL 61801), and James F. Zachary (Univ. of Illinois, Urbana, IL 61801)

Three kinds of solid tumors were acquired and scanned *in vivo* ultrasonically. The first tumor series (fibroadenoma) was acquired from tumors that had spontaneously developed in rats. The second tumor series was acquired by culturing a carcinoma cell line (4T1-MMT) in culture media and injecting the cells into Balb/c mice. The third tumor was acquired by transplanting a soft-tissue sarcoma cell line (EHS) into C57BL mice. The tumors were allowed to grow to 1 cm in size and then scanned ultrasonically. The scatterer properties of average scatterer diameter and acoustic concentration were estimated using a Gaussian form factor from the backscattered ultrasound measured from the tumors. Parametric images of the tumors were constructed utilizing estimated scatterer properties for regions of interest inside the tumors. The parametric images showed distinct differences between the various tumor types. Quantitatively, the tumors could be distinguished through feature analysis plots of average scatterer size versus acoustic concentration. Comparison with photomicrographs of the tumors showed structures similar in size to the ultrasound estimates. [Work supported by NIH Grant F32 CA96419 to MLO and by the University of Illinois Research Board.]

8:45

5aBB4. Evaluation of the DORT method for the detection of microcalcifications in the breast. Jean-Luc Robert, Claude Cohen-Bacrie (Philips Res. USA, 345 Scarborough Rd., Briarcliff Manor, NY 10510-2099), Claire Prada, and Mathias Fink (Universite Denis Diderot, 75231 Paris Cedex 05, France)

The DORT method (French acronym for diagonalization of the time reversal operator) is derived from the theory of iterative time reversal mirroring. It consists of a singular value decomposition of the time reversal operator obtained through single element transmissions and receptions. The number of eigenvalues relates to the number of bright point scatterers in the medium, and each eigenvector is the transmit signals that focuses on each bright scatterer. However, the signal-to-noise ratio (SNR) resulting from a single element transmit is low, which negatively impacts the sensitivity of the DORT method. This work consists of an adaptation of the DORT method to an imaging mode. Focused transmissions in the medium are used and a windowing preprocessing operation on the received signals significantly increases the sensitivity. The more robust behavior of this modified DORT method is tested on Field II simulated data and then on a phantom made of strings of different material embedded in speckle. Data on freshly excised breast surgical samples containing microcalcifications were processed and every microcalcifications was detected. Finally, results on *in vivo* acquisitions prove that the technique can accurately detect and locate microcalcifications in nonhomogeneous media with aberration.

9:00

5aBB5. Scatterer size and concentration estimation technique based on a 3D acoustic impedance map from histologic sections. Jonathan Mamou, Michael L. Oelze, William D. O'Brien, Jr. (Dept. of Elec. and Computer Eng., Univ. of Illinois, 1406 W. Green St., Urbana, IL 61801), and James F. Zachary (Univ. of Illinois, Urbana, IL 61801)

Accurate estimates of scatterer parameters (size and acoustic concentration) are beneficial adjuncts to characterize disease from ultrasonic backscatterer measurements. An estimation technique was developed to obtain parameter estimates from the Fourier transform of the spatial autocorrelation function (SAF). A 3D impedance map (3DZM) is used to obtain the SAF of tissue. 3DZMs are obtained by aligning digitized light microscope images from histologic preparations of tissue. Estimates were obtained for simulated 3DZMs containing spherical scatterers randomly located: relative errors were less than 3%. Estimates were also obtained from a rat fibroadenoma and a 4T1 mouse mammary tumor (MMT). Tissues were fixed (10% neutral-buffered formalin), embedded in paraffin, serially sectioned and stained with H&E. 3DZM results were compared to estimates obtained independently against ultrasonic backscatter measurements. For the fibroadenoma and MMT, average scatterer diameters were 91 and 31.5 μm, respectively. Ultrasonic measurements yielded average scatterer diameters of 105 and 30 μm, respectively. The 3DZM estimation scheme showed results similar to those obtained by the independent ultrasonic measurements. The 3D impedance maps show promise as a powerful tool to characterize ultrasonic scattering sites of tissue. [Work supported by the University of Illinois Research Board.]

9:15

5aBB6. Ultrasonic tissue characterization for the classification of prostate tissue. Ulrich Scheipers, Helmut Ermert (Lehrstuhl fuer Hochfrequenztechnik, Ruhr-Universitaet Bochum [RUB], Germany, ulrich.scheipers@rub.de), Katharina Koenig, Hans-Joerg Sommerfeld, Miguel Garcia-Schuermann, Theodor Senge (Urologische Universitaetsklinik der RUB, Marienhospital Herne, Germany), and Stathis Philippou (Institut für Pathologie, Augusta-Krankenanstalt Bochum, Germany)

Radio-frequency ultrasound echo data of the prostate are captured during routine examinations with standard ultrasound equipment. The data are directly transmitted to a PC and subdivided into numerous regions of interest. Several parameters describing the histological characteristics of the underlying tissue are calculated from the frequency spectrum and from the demodulated signal of the underlying echo data. Parameters are fed into two adaptive network-based fuzzy inference systems working in parallel. One system is used to classify hypo- and hyperechoic tumors, the other system is used to detect isoechoic tumors within the normal prostate tissue. Subsequent morphological analysis combines clusters to mark areas of similar tissue characteristics. Classification results are presented as two-dimensional malignancy maps and as volumetric reconstructions of the whole organ. Radio-frequency ultrasonic echo data of 100 patients have been recorded. Tissue samples following radical prostatectomies are used as the gold standard. The area under the ROC curve is $A = 0.86 \pm 0.01$ for hypo- and hyperechoic tumors and $A = 0.84 \pm 0.02$ for isoechoic tumors using leave-one-out cross validation over patient datasets.

9:30

5aBB7. Sonoporation of cell using therapeutic ultrasound for drug and gene delivery. Hua Pan, Fred Sieling, Yun Zhou, Hesheng Wang, Jianmin Cui, and Cheri Deng (Dept. of Biomed. Eng., Case Western Reserve Univ., 10900 Euclid Ave., Cleveland, OH 44106-7207)

Recent studies of ultrasound methods for targeted drug delivery and nonviral gene transfection revealed new, advantageous possibilities. These studies utilized ultrasound contrast agents, commonly stabilized microbubbles, to facilitate delivery and suggested that ultrasound delivery resulted from cell sonoporation, the formation of temporary pores in the cell membrane induced by ultrasound. In this study, voltage clamp techniques were used to obtain real-time measurements of sonoporation of single Xenopus oocytes in the presence of Optison. Ultrasound increased the transmembrane current as a direct result of decreased membrane resistance due to pore formation. We observed a distinct delay of sonoporation following ultrasound activation and characteristic stepwise increases of transmembrane current throughout ultrasound duration. We discovered that the resealing of the cell membrane following ultrasound exposure required calcium entering the cell through ultrasound-induced pores.

9:45

5aBB8. A comparison study in cell transfection: Which one is better, sonoporation versus electroporation? Junru Wu, Jason Pepe (Dept of Phys., Univ. of Vermont, Burlington, VT 05405), and Mercedes Rincon (Univ. of Vermont, Burlington, VT 05405)

An experimental study has been performed for cell suspensions to compare efficiency of cell transfection, which is the process of introducing recombinant DNA into eukaryotic cells (eukaryotic cells have chromosomes with nucleosomal structure) and subsequently integrating that DNA into the recipient cell's chromosomal DNA. It was demonstrated that electroporation was superior to sonoporation in terms of viability (65.8 2.3% vs. 50.8 4.15%) and transfection efficiency (15.83 3.5% vs. 7.53 0.4%) for Jurkat lymphocytes (nonprimary cells), and sonoporation was better in terms of viability (64.8 1.51% vs. 53.7 1.53 %) and transfection efficiency (2.73 0.21% vs. 0.43 0.06%) for human peripheral blood mononuclear cells (primary cells). The electroporation was performed using a Gene Pulser II Apparatus with voltage of 250 V, and the sonoporation was achieved using 2-MHz pulsed ultrasound exposure (ISPPA$=$80 W/cm^2) assisted with encapsulated bubbles (Optison).

10:00

5aBB9. HIFU-induced gene activation *in vitro*. Yunbo Liu, Pei Zhong (Dept. of Mech. Eng., Duke Univ., Durham, NC 27708), Takashi Kon, and Chuanyuan Li (Duke Univ., Durham, NC 27708)

This work investigated the inducible gene activation in cancer cells that were sublethally injured during HIFU treatment. HeLa cells were transfected by an adenovirus vector that encodes GFP under the control of hsp70B promoter, leading to about 65% transfection efficiency. A volume of 10 μL transfected HeLa cells in suspension (5×10^7 cells/ml) were placed at the bottom of a PCR tube so that the cell suspension could be heated to a peak temperature of 50 °C, 60 °C, and 70 °C for 120, 10, and 1 s, respectively, by a focused 1.1-MHz HIFU transducer operated at a peak negative pressure of -2.7 MPa at different duty cycles. One day after HIFU treatment, cell viability was determined to be 63%, 35%, and 18%, respectively, based on Trypan Blue exclusion test. Importantly, in all test groups, inducible GFP expression was detected in about 40%–50% of the surviving cells with GFP intensity increased by 25-fold based on flow cytometry analysis. These results demonstrate that even under the short exposure duration of HIFU treatment, inducible gene expression could be produced in sublethally injured cell population *in vitro*. Further studies are underway to explore the optimal HIFU condition for gene activation *in vivo*.

10:15

5aBB10. Why does humming clean your maxillar sinuses of NO gas? Johan Sundberg, Svante Granqvist (Speech Music Hearing, KTH, SE-100 44 Stockholm, Sweden, pjohan@speech.kth.se), Eddie Weitzberg, and Jon Lundberg (Karolinska Inst., Karolinska SE-17177, Sweden)

Recent measurements have shown that the nitric oxide (NO) produced in the maxillar sinuses can be evacuated by producing nasal murmur; the NO content of the nasal airflow increases under these conditions (Maniscalco *et al.*, 2003). Experiments were carried out to test the hypothesis that this effect is caused by resonance in the nasal tract and sinus cavities. A model was constructed where an airflow from a pressure tank was modulated at different frequencies. This airstream was passed through a tube with a radial hole constituting the neck of a Helmholtz resonator with a gas containing NO. The NO content of the air streaming out of the tube was measured. This NO content varied when the location of the resonator, its air volume, or the modulation frequency of the airflow was changed. The relevance of these three factors was also tested by means of a computer model of the system. The significance of the modulation frequency was also analyzed in human subjects by injecting a modulated airflow through one nostril and measuring NO content of the airflow exiting the other nostril, the results showing a dependence on the modulation frequency that differed considerably between subjects [Maniscalco *et al.*, "Assessment of nasal and sinus nitric oxide output using single-breath humming exhalations," Eur. Respir. J. **22**, 323–329 (2003)].

FRIDAY MORNING, 28 MAY 2004 LIBERTY 5, 8:25 TO 11:15 A.M.

Session 5aEA

Engineering Acoustics: Computational Acoustics, Ultrasonics and Applications

Elizabeth A. McLaughlin, Cochair
Naval Undersea Warfare Center, 1176 Howell Street, Newport, Rhode Island 02841

Juan Arvelo, Cochair
Applied Physics Laboratory, Johns Hopkins University, 11100 Johns Hopkins Road, Laurel, Maryland 20723-6099

Chair's Introduction—8:25

Contributed Papers

8:30

5aEA1. A Galerkin discontinuous method for CAA: The reasons of a choice. Philippe Delorme and Christophe Peyret (ONERA, BP 72, 92322 Chatillon, France)

As for many others disciplines, the numerical simulation is becoming a powerful tool for the study of the propagation of small disturbances in a heterogeneous flow (aero-acoustics). Some main applications are, on the one hand, aeronautics for the noise of the aircraft and, on the other hand, the propagation of the sound in the atmosphere. In general the development of a computer code goes by four phases: physical modeling, mathematic analysis, strategy of discretization, programming and validation. At each of these stages some specific choices are carried out. For physical modeling, it is necessary to choose a representative model. Then, it should be checked that the problem is well posed in existence and unicity. For the strategy of discretization the choice resides between the different methods and is guided by the geometries to study. Finally, the choice to use or not the possibility of strongly paralleling influences programming of the code. In this paper the reasons are described which led us to use a method of discontinuous the Galerkin type for the CAA. This method is then presented as well as architecture of the code. Finally, all this is illustrated for examples of simulations.

8:45

5aEA2. Approaches to simulate electromechanical coupling in voided piezoelectric materials. Juan Arvelo, Jr. (Johns Hopkins Univ., Appl. Phys. Lab., 11100 Johns Hopkins Rd., Laurel, MD 20723-6099) and Ilene Busch-Vishniac (Johns Hopkins Univ., Baltimore, MD 21218-2681)

Recent work has demonstrated piezoelectric behavior in voided polymers. It is thought that the piezoelectricity results from the creation of space charges bordering the voids when the material is exposed to an electric field in excess of the breakdown strength. The resulting voided piezoelectric material is more flexible than conventional piezoelectric ceramics, and may be far less expensive to manufacture. Accurate models of piezoelectricity in voided materials are needed to explore the limits of performance and their use under a wide range of conditions. We are using finite element analysis to develop a model of the mechanical, electrical, and piezoelectric behavior in voided materials. Results for some of the materials of current interest, such as low-density polypropylene (LDPP) and porous polytetrafluoroethylene (PTFE) will be presented. [Work performed under a JHU/APL sabbatical to the JHU/WSE.]

9:00

5aEA3. Simulation of linear aeracoustic propagation in lined ducts with discontinuous Galerkin method. Christophe Peyret and Philippe Delorme (ONERA BP, 72 Chatillon, France)

The simulation of the acoustic propagation inside a lined duct with nonuniform flow still presents problems when geometry is complex. To handle computations on a complex geometry, without consequential effort, unstructured meshes are required. Assuming irrotational flow and acoustic perturbation, a well-posed finite element method based on the potential equation is established. But, the effect of the thin boundary layer is then neglected, which is not relevant to the acoustical processes occurring near the lining. Recent works have focused on the tremendous interest of the Galerkin discontinuous method (GDM) to solve Euler's linearized equations. The GDM can handle computations on unstructured meshes and introduces low numerical dissipation. Very recent mathematical works have established, for the GDM, a well-posed boundary condition to simulate the lining effect. Results computed with the GDM are presented for a uniform cross section lined duct with a shear flow and are found to be in good agreement with the modal analysis results, thereby validating the boundary condition. To illustrate the flexibility of the method other applications dealing with instabilities, air-wing diffraction, and atmospheric propagation are also presented.

9:15

5aEA4. Engineering acoustic lenses with help from evolution. Andreas Håkansson, José Sánchez-Dehesa, and Lorenzo Sánchis (Nanophotonic Technol. Ctr. and Dept. of Electron. Eng., Polytechnic Univ. of Valencia, Spain)

Optimization engineering through evolutionary algorithms have proven to be very efficient, especially in hard problems containing a large set of optimization parameters. Like evolution this family of algorithms is able to tackle enormous complex problems with fairly simple means. Here, a simple genetic algorithm [J. H. Holland, *Adaptation in Natural and Artificial Systems* (Univ. of Michigan, Ann Arbor, 1975)] is used in conjunction with the multiple scattering theory [L. Sánchis *et al.*, Phys. Rev. B **67**, 035422 (2003)] to fabricate a new generation of acoustic devices based on a discrete number of cylindrical scatterers. In particular, acoustic lenses [F. Cervera *et al.*, Phys. Rev. Lett. **88**, 023902 (2002)] with flat surfaces have been designed to focus the sound in a fixed focal point for one or multiple frequencies. Each scatterer is carefully placed using the optimization method within the preset boundary conditions, to maximize the pressure contribution in the chosen focal spot. With this method acoustic lenses with very low f-numbers of the order 0.3 and with amplifications over 12 dB have been estimated using a reduced number of scatterers ($\sim$60). Preliminary results obtained from the experimental realization of the designed devices confirm our predictions.

9:30

5aEA5. An application of boundary element method calculations to hearing aid systems: The influence of the human head. Karsten B. Rasmussen (Oticon A/S, Strandvejen 58, DK-2900 Hellerup, Denmark) and Peter Juhl (Univ. of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark)

Boundary element method (BEM) calculations are used for the purpose of predicting the acoustic influence of the human head in two cases. In the first case the sound source is the mouth and in the second case the sound is plane waves arriving from different directions in the horizontal plane. In both cases the sound field is studied in relation to two positions above the right ear being representative of hearing aid microphone positions. Both cases are relevant for hearing aid development. The calculations are based upon a direct BEM implementation in Matlab. The meshing is based on the original geometrical data files describing the B&K Head and Torso Simulator 4128 combined with a 3D scan of the pinna.

9:45

5aEA6. Dissipative silencers with an extended inlet/outlet and baffles. Ahmet Selamet, Iljae Lee, Mubing Xu (The Ohio State Univ., 930 Kinnear Rd., Columbus, OH 43210, selamet.1@osu.edu), and Norman Huff (Owens Corning Automotive, Novi, MI 48377)

The acoustic characteristics of a single-pass perforated dissipative silencer were investigated experimentally and numerically by Selamet *et al.* [J. Acoust. Soc. Am. **109**, 2364 (2001)]. The current study extends this work by considering variations in the internal structure of the dissipative silencer. In addition to the boundary element method (BEM) introduced earlier, a multi-dimensional analytical approach is now developed to investigate the wave modes and transmission loss. Both methods are then employed to study the effect of an extended inlet and outlet on the acoustic behavior of the silencer. BEM is further used to explore the effect of baffles and air space inside the dissipative chamber. The location and number of baffles inside the dissipative chamber are shown to have a significant influence on the transmission loss.

10:00

5aEA7. Acoustic waves in discrete media, similarities at meso and nano scales. Hasson Tavossi (Dept. of Physical and Environ. Sci., Mesa State College, School of Math. and Physical Sci., 1100 North Ave., Grand Junction, CO 81501)

In this paper the similarities between the acoustic behavior of a discrete medium of random arrangement of solid spheres, at mesoscopic scales, and thermal vibrations of lattice ions in crystalline solid, at nano scales, are investigated. Results for the ultrasonic waves in random media show that such effects as cutoff frequency, wave dispersion, energy distribution in vibration modes, wave attenuation by scattering and absorption, observed in discrete media, have close resemblance to the similar phenomena observed at atomic scales, such as phonons or lattice thermal vibration of atoms in the crystalline solids. For example, the cutoff frequency in lattice vibration is related to the interatomic spacing, similarly, the cutoff frequency at mesoscopic scales depends on grain separation or grain size. These similarities are also observed in wave scattering and attenuation and their dependence on wave number, kR ($k=2\pi/\lambda$ and R is particle size). In this paper experimental measurements with data analysis on cutoff frequency, wave attenuation, and wave dispersion related to the above mentioned similarities will be presented, and the extension of these findings to the behavior of acoustic waves in other discrete media will be discussed.

10:15

5aEA8. Ultrasonic imaging in noisy environment. John Lewis, Matt Kaiser, James Irwin, Jr., and Jose Sanchez (Dept. of Elec. & Computer Eng., Bradley Univ., 1501 W. Bradley Ave., Peoria, IL 61625, jhirwin@bradley.edu)

Often times fossils are found buried within rocks of similar composition to the fossil. Ultrasonic imaging is becoming a popular form of nondestructive testing. The presentation covers techniques for imaging a fossil through nondestructive testing in three dimensions without risking damage to the fossil by removing extraneous material. The lecture begins with an illustration of immersion testing that uses longitudinal transducers to construct images of objects buried within a noisy environment of similar composition. Preliminary results give the location and shape of an object buried under a homogeneous layer. The first step to attaining this goal is to merely locate an object buried in a noisy debris field. A single sensor is all that is required to locate an object; however, two sensors allow the object's position to be triangulated. Techniques for using single and multiple sensors are investigated to improve the search and imaging algorithms. The lecture continues with an in-depth discussion of the MATLAB signal-processing algorithms, which are required to generate a visual representation of the object. With some modification this technique also promises to be useful for security screening and landmine detection.

10:30

5aEA9. Time reversal interactive objects. Ros Ki Ing (Sensitive Object, Res. and Development, 10 rue Vauquelin, 75005 Paris, France), Nicolas Quieffin, Stefan Catheline, and Mathias Fink (Paris 7 Univ., 75231 Cedex, Paris)

Time reversal has shown to be a fruitful concept in nondestructive testing in underwater acoustic or in ultrasonic imaging. In this paper this technique is adapted in the audible range to transform every day objects into tactile sensitive interfaces. A quick historical background is presented in the ultrasonic field and specially in chaotic cavity. In all time reversal experiments, it is demonstrated that a wave field spatially and temporally recorded is able to back propagate to its source. In other words, the field contains all the information on the location of the source. In the interactive experiments, it is shown that touching an object like a window, a table or a world globe generates an acoustic field easily detectable with one or two acoustic sensors. Using the concept of time reversal, the source location is deduced in real time. Then, touching objects at specific locations (virtual switches) is used to activate devices. Such devices are for example lights, stereo volume, or computer software. From a technical point of view, all these interactive experiments just use some computation easily performed with a standard personnel computer.

10:45

5aEA10. High-sensitivity photoacoustic leak testing. Eric Huang, David R. Dowling (Dept. of Mech. Eng., Univ. of Michigan, Ann Arbor, MI 48109-2133), Timothy Whelan (Honeywell FM&T LCC, Kansas City, MO 64141-6159), and John L. Spiesberger (Univ. of Pennsylvania, Philadelphia, PA 19104-6316)

The photoacoustic effect may be exploited for detection and localization of gas leaks on the surface of otherwise sealed components. The technique involves filling the test component with a photoactive tracer gas, and irradiating the component to produce photoacoustic sound from any leak site where a tracer gas cloud forms. This presentation describes demonstration experiments utilizing 10.6-μm radiation from a nominally 145-W carbon-dioxide laser with sulfur hexafluoride as a tracer gas. Here, photoacoustic sounds from six NIST-traceable calibrated leak sources with leak rates between 1 cc in 4.6 h, and 1 cc in 6.3 years were recorded with 12 microphones in a bandwidth from 3 to 80 kHz. Bartlett matched-field processing of the microphone array measurements both detect and localize these leaks when the leak and the array are separated by 152 mm. These experiments suggest that the sensitivity of photoacoustic leak testing may reach or even exceed the capabilities of the most sensitive commercial leak test systems using helium mass-spectrometers. Comparison of the measured results and an analytical scaling law suggests that tracer cloud geometry influences the photoacoustic signal amplitude. [Work supported by the U.S. Dept. of Energy.]

11:00

5aEA11. Evaluation of electromagnetic field interference on sound-pressure level measurements. Luigi Maxmilian Caligiuri and Adolfo Sabato (Dept. of Mech., Faculty of Eng., Univ. of Calabria—Via P. Bucci 87030 Arcavacata di Rende [CS], Italy, maxmc@tiscali.it)

The recently published IEC 61672-1 defines new standards for sound-level meters and calibrators, giving performance specifications, environmental and electromagnetic criteria, in order to overcome some important faults of the previous IEC 651-1979 and IEC 804-1985. Both of these standards have been in fact amended to include, in particular, specifications for immunity of electrostatic and electromagnetic fields. The interference of such nonacoustics factors can represent, if nonadequately recognized and quantified, an important uncertainty source in sound-pressure level measurements, able to cause errors on measurement result, variable between fractions of dB and some dB. In the present paper we will show the application of a measurement methodology able to quantify, within the framework of IEC 61672-1, the influence of electromagnetic fields on environmental noise measurements, carried out by a class I (according to IEC 651 and 804) integrating sound-level meter. In this way, we have analyzed the effect of different typologies of interfering field sources (characterized by different spectrum and emission features) on some very frequent noise-measurement configuration.

Session 5aMU

Musical Acoustics and Psychological and Physiological Acoustics: Neurophysiology of Playing a Musical Instrument

Ingo R. Titze, Chair
Department of Speech Pathology and Audiology, University of Iowa, 330 WJSHC, Iowa City, Iowa 52242-1012

Chair's Introduction—9:00

Invited Papers

9:05

5aMU1. Movement amplitude and tempo change in piano performance. Caroline Palmer (Dept. of Psychol., McGill Univ., 1205 Dr. Penfield Ave., Montreal, QC H3A 1B1, Canada, caroline.palmer@mcgill.ca) and Simone Dalla Bella (Kazimierz Wielki Univ., Bydgoszcz, 85–867 Poland)

Music performance places stringent temporal and cognitive demands on individuals that should yield large speed/accuracy tradeoffs. Skilled piano performance, however, shows consistently high accuracy across a wide variety of rates. Movement amplitude may affect the speed/accuracy tradeoff, so that high accuracy can be obtained even at very fast tempi. The contribution of movement amplitude changes in rate (tempo) is investigated with motion capture. Cameras recorded pianists with passive markers on hands and fingers, who performed on an electronic (MIDI) keyboard. Pianists performed short melodies at faster and faster tempi until they made errors (altering the speed/accuracy function). Variability of finger movements in the three motion planes indicated most change in the plane perpendicular to the keyboard across tempi. Surprisingly, peak amplitudes of motion before striking the keys increased as tempo increased. Increased movement amplitudes at faster rates may reduce or compensate for speed/accuracy tradeoffs. [Work supported by Canada Research Chairs program, HIMH R01 45764.]

9:25

5aMU2. Physiology, anatomy, and plasticity of the cerebral cortex in relation to musical instrument performance. Mark Jude Tramo (Dept. of Neurol., Harvard Med. School and Massachusetts General Hospital and the Inst. for Music and Brain Sci., Boston, MA 02114-2696, mtramo@hms.harvard.edu)

The acquisition and maintenance of fine-motor skills underlying musical instrument performance rely on the development, integration, and plasticity of neural systems localized within specific subregions of the cerebral cortex. Cortical representations of a motor sequence, such as a sequence of finger movements along the keys of a saxophone, take shape before the figure sequence occurs. The temporal pattern and spatial coordinates are computed by networks of neurons before and during the movements. When a finger sequence is practiced over and over, performance gets faster and more accurate, probably because cortical neurons generating the sequence increase in spatial extent, their electrical discharges become more synchronous, or both. By combining experimental methods such as single- and multi-neuron recordings, focal stimulation, microanatomical tracers, gross morphometry, evoked potentials, and functional imaging in humans and nonhuman primates, neuroscientists are gaining insights into the cortical physiology, anatomy, and plasticity of musical instrument performance.

9:45

5aMU3. Speed, accuracy, and stability of laryngeal movement in singing. Ingo R. Titze (Dept. of Speech Pathol. and Audiol., Univ. of Iowa, Iowa City, IA 52242 and Natl. Ctr. for Voice and Speech, Denver Ctr. for the Performing Arts, Denver, CO 80204, ititze@dcpa.org)

Motor performance is often quantified in terms of speed, strength, accuracy, and stability of a target gesture, or maintaining a given posture. In the vocal system, this involves primarily the intrinsic laryngeal muscles and the respiratory muscles. Agonist–antagonist pairs of muscles are used to position the vocal folds for phonation (vocal onset), for pitch change, and for registration (as in yodeling). Maximum speed and accuracy are discussed for vocal embellishments such as trills, trillo, scales, arpeggios, yodel, and glissando. This speed and accuracy are interpreted in terms of muscle twitch and tetanic responses obtained *in vitro* on animal muscles, from electromyographic recordings on humans, and from muscles not easily tested on humans. The laryngeal reflex system is also described, particularly with regard to its ability to stabilize (or destabilize) neurologic tremor originating from the central nervous system.

Contributed Paper

10:05

5aMU4. Flute "breath support" perception and its acoustical correlates. Isabelle A. Cossette and Patrick Sabourin (McGill Univ., Montreal, QC, Canada)

Music educators and performers commonly refer to "breath support" in flute playing, yet the term "support" is neither well-defined nor consistently used. Different breathing strategies used by professional flautists who were instructed to play with and without support were previously identified by the authors. In the current study, 14 musical excerpts with and without support were recorded by five professional flautists. Eleven professional flautists listened to the recordings in a random order and ranked (1 to 6) how much of the following sound qualities they judged to be in each example: support, intonation, control and musical expressiveness. Answers to the test showed that musical expressiveness was associated more closely with the supported excerpts than the answers about support itself. The ratings for each sound quality were highly intercorrelated. Acoustical parameters were analyzed (frequency and centroid variation within each note) and compared with the results of the perception test in order to better understand how the acoustical and psychological variables were related. The acoustical analysis of the central part of the notes did not show evident correlation with the answers of the perception test. [Work funded by the Social Sciences and Humanities Research Council of Canada.]

FRIDAY MORNING, 28 MAY 2004 CONFERENCE ROOM D, 7:50 A.M. TO 12:00 NOON

Session 5aNS

Noise: Noise in Large Cities I

Daniel R. Raichel, Chair
2727 Moore Lane, Fort Collins, Colorado 80526

Chair's Introduction—7:50

Invited Papers

8:00

5aNS1. Noise and soundscape in Rome. Giovanni Brambilla (CNR-Inst. of Acoust., via del Fosso del Cavaliere 100, 00133 Rome, Italy)

Noise pollution is an old problem in Rome. In 45 B.C. the Lex Julia Municipalis limited carriage traffic in the city center to specific times. Road traffic constitutes the most important and widespread noise source, and several investigations have been conducted since 1972, some aimed at developing a numerical model for predicting the hourly LAeq level. In order to reduce the large impact of this type of noise some measures have been carried out, including surfacing with porous asphalt, erection of noise barriers, limitation in time and spacing of private traffic, etc. However, most of the public complaints deal with noise from equipment operation and recreational activities rather than transportation systems. Moreover, the most famous tourist areas opened to pedestrians only are not as quiet as expected but their sound environment is usually rated more acceptable than noise from other sources at the same level. In compliance with the Italian legislation on noise, the Municipality of Rome issued a noise zoning code for its own territory, and a noise mapping is in progress, pursuant to the requirements of the 2002/49/EC European directive. A Geographical Information System has been also developed to manage all the aspects of noise pollution.

8:20

5aNS2. Environmental noise in Beijing. Jing Tian (Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC) and Yi Wang (Beijing Ctr. of Environ. Protection and Monitoring, Beijing 100000, PROC)

Beijing is a city under rapid development. More than 60% of the complaints on the environment comes from noise pollution. With a road system of total length of 4200 km and thousands of construction sites, road traffic, and construction are the most important noise sources in Beijing. One third of its population of nearly 14 million inhabitants and more than 3 million visitors is influenced by traffic noise and 16% of its inhabitants are living along various roads and expressways. A series of noise control measures have been taken since the early 1980s, such as road system reconstruction and improvement, industrial noise source control, vehicle noise control, as well as various administrative steps, which reduced the traffic noise by 7 dBA with the number and average flux of vehicles multiplied by more than three times and kept the city in a day–night average level around 64 dBA. In this paper, the environmental noise legislation in Beijing is briefly introduced. The historical progress in noise control is reviewed. The noise pollution up-to-date is given and discussed. Some possible improvements and potential opportunities are presented.

8:40

5aNS3. The meaning of city noises: Investigating sound quality in Paris (France). Daniele Dubois, Catherine Guastavino, Valerie Maffiolo (CNRS LCPE/LAM 11 rue de Lourmel (F) 75015 Paris, France, ddubois@ccr.jussieu.frieu.fr), Catherine Guastavino (McGill Univ., Montreal, Canada), and Valerie Maffiolo (FTR&D, Lannion, France)

The sound quality of Paris (France) was investigated by using field inquiries in actual environments (open questionnaires) and using recordings under laboratory conditions (free-sorting tasks). Cognitive categories of soundscapes were inferred by means of psycholinguistic analyses of verbal data and of mathematical analyses of similarity judgments. Results show that auditory judgments mainly rely on source identification. The appraisal of urban noise therefore depends on the qualitative evaluation of noise sources. The salience of human sounds in public spaces has been demonstrated, in relation to pleasantness judgments: soundscapes with human presence tend to be perceived as more pleasant than soundscapes consisting solely of mechanical sounds. Furthermore, human sounds are qualitatively processed as indicators of human outdoor activities, such as open markets, pedestrian areas, and sidewalk cafe districts that reflect city life. In contrast, mechanical noises (mainly traffic noise) are commonly described in terms of physical properties (temporal structure, intensity) of a permanent background noise that also characterizes urban areas. This connotes considering both quantitative and qualitative descriptions to account for the diversity of cognitive interpretations of urban soundscapes, since subjective evaluations depend both on the meaning attributed to noise sources and on inherent properties of the acoustic signal.

9:00

5aNS4. The sound of Berlin. The noise annoyance you love to hate. Brigitte Schulte-Fortkamp and Cay Hehner (Inst. of Tech. Acoust., TU-Berlin, Einsteinufer 25, D-10587 Berlin, Germany, brigitte.schulte-fortkamp@tu-berlin.de)

Noise is part of life in Berlin. The Berlin Sound is often described as Berlin is a loud city, but this sound is also a matter of identification. Just like human beings, metropolitan areas may be identified by their gait. People enjoy living in Berlin but they also claim negative health effects and often permanent annoyance caused by daily environmental noise. For instance 1 119 714 cars demonstrate the volume of traffic from 6.30–9.00 a.m, and another impressive descriptor of the volume of traffic is the number of 1 567 600 cars from 4.00–6.30 p.m. In a representative survey of 2000 adults, almost all German (80%) citizens are affected by some level of noise pollution. The predominant source of noise in residential areas is road traffic, which remains a nuisance for over half the population, and a source of serious annoyance for some 18%. Next to road traffic, air traffic is the most important transport-related source of annoyance, followed by rail traffic noise. Results from a qualitative survey in a residential area of Berlin give insights into the different perspectives of the noise perception of the investigated subjects and of important parameters with respect to daily life.

9:20

5aNS5. Noise and social deprivation in an urban environment. Bridget Shield (Dept. of Eng., London South Bank Univ., London SE1 0AA, UK, shieldbm@lsbu.ac.uk) and Julie Dockrell (London Univ., London W1H 0AA, UK)

Noise levels have been measured inside and outside approximately 170 schools in London, England, as part of a project to investigate the effects of noise on the cognitive performance and academic attainments of children of primary school age in London. As well as providing data on individual schools the survey has provided a portrait of the noise climate across London. In addition to noise levels, the external noise sources present at each measurement location have been identified. The locations in which external schools noise levels were measured include areas where road traffic is the predominant noise source and areas near Heathrow Airport where aircraft are the major source. In addition to noise data, the following socioeconomic information has been obtained for a majority of the schools: percentages of children at each school having free school meals and the numbers for whom English is not the first language. The first of these is known to be a reliable indicator of social deprivation in an area. The relationship between noise and these socioeconomic factors has been established, which shows that, as might be expected, the higher noise levels in London are associated with the areas of greater social deprivation.

9:40

5aNS6. Noise in Mexico City. Sergio Beristain (IMA, ESIME, IPN. P.O. Box 12-1022, Narvarte, 03001, Mexico D. F., Mexico, sberista@hotmail.com)

Mexico City is known to be the largest city in the world, inhabited by some 20 percent of the national population, so noise pollution is not strange to it, particularly in view of the fact that industry is not concentrated, but rather spread throughout the city. The international airport also lies within the city limits, in the midst of residential areas. The heavy traffic during rush hours in the morning and in the evening and the activities of the populace, together with special events, produce a noise problem that is difficult to assess and to solve. Nevertheless, with educational programs begun several years ago and noise campaigns planned for the near future, in addition to existing regulations, the problem is not completely out of control. This paper presents a discussion of the general noise problem and describes how authorities and institutions are dealing with it.

10:00–10:20 Break

10:20

5aNS7. Noise in large cities in Brazil. Samir N. Y. Gerges (Mech. Eng. Dept. (EMC), Federal Univ. of Santa Catarina (UFSC), Florianopolis, SC, Brazil)

Large cities' noise is considered by the World Health Organization to be the third most hazardous pollution, preceded by air and water pollution. In urban centers, in general, and especially in developing countries such as Brazil, large populations are affected by excessive noise due mainly to traffic flow. The Brazilian Federal Government specifies noise limits, but each state can enforce its own set of noise limits, providing they are lower. The rapid economic growth, together with large migration of northern Brazilians to the

developing southern urban areas in search of more lucrative jobs in construction and industrial sectors, resulted in a fast increase in activities such as vehicle and bus traffic, home construction, and development of all necessary infrastructures to support this growth. Urban noise in Brazil has been receiving the attention of national authorities only since 1990, when the Federal Government approved the first ''Program of Community Silence,'' based on ISO R 1996-1971. This paper highlights the noise situation in the five largest and most populated cities in Brazil: Sao Paulo, Rio de Janeiro, Belo Horizonte, Porto Alegre and Curitiba [Zannin *et al.*, Appl. Acoust. **63**, 351–358 (2002)].

10:40

5aNS8. Noise issues in Kanagawa Prefecture. Shigenori Yokoshima (Kanagawa Environ. Res. Ctr. 1-3-39, Shinomiya, Hiratsuka, Kanagawa Prefecture 254-0014, Japan, yokoshima@k-erc.pRef. kanagawa.jp) and Akihiro Tamura (Yokohama Natl. Univ., Hodogaya, Yokohama 240-8501, Japan)

In Kanagawa Prefecture, bordering Tokyo Metropolis and the third most densely populated prefecture in Japan, various noises have caused serious problems in terms of living environment preservation and human health protection. This paper describes present states of noise issues in Kanagawa. Road traffic noise, remaining one of serious pollution issues, was monitored at a total 217 sites along trunk roads in Kanagawa from fiscal year 2000 to 2002. The percentage of the sites that achieve environmental quality standards for road traffic noise was approximately 20%. Noise caused by Tokaido Shinkansen trains, of which the total daily number is 287, also has negative impacts on inhabitants along the railway. As a result of the noise measurement from fiscal year 1994 to 2002, about 80% of the measurement sites exceeded environmental quality standards for Shinkansen railway noise during the years. In the areas surrounding the Atsugi Base, noise generated by training flights damagingly affects inhabitants' daily life. The number of complaints due to the noise was largest among noise issues. Moreover, neighborhood noises, noises emitted during the nighttime operation of bars, restaurants and shops, and noises produced by work in out-of-door yards have recently provoked social issues.

11:00

5aNS9. Managing environmental noise in Hong Kong. Kai Ming Li (Dept. of Mech. Eng., The Hong Kong Polytech. Univ., Hung Hom, Hong Kong)

Hong Kong is well known for its economic vibrancy and its hyper densely population: more than 7 million people living in a total area of slightly over 1000 square kilometers of hilly areas. Most of these people live and work in about 20% of the total land area, resulting in probably the highest densities in the world. The high population density is also matched by a large number of vehicles running in the roads. At present, there are over 400 000 vehicles operating on a highway network less than 1000 km in length. With all these factors plus many urban activities associated with the rapid growth and development, noise is an important environmental issue in the city. Although there are many dimensions for the quality of life, the acoustic environment is undoubtedly an essential part of it, especially when the people's aspirations increase as the society develops. This paper summarizes the development of strategies for controlling environmental noise in Hong Kong in the past two decades. The current situation will be addressed and a proposal for an improved traffic noise policy will be presented. [Work supported by the Research Grants Council, and The Hong Kong Polytechnic University.]

11:20

5aNS10. Noise enforcement in cities. Eric Zwerling (Rutgers Univ. Noise Tech. Assistance Ctr., Dept. of Environ. Sci., 14 College Farm Rd., New Brunswick, NJ 08901)

Noise enforcement programs (NEPs) in cities face a number of unique obstacles, including the necessity to conduct sound level measurements in a complex acoustical environment. It is important to regularly and objectively review the efficacy of the program in totality. A successful NEP consists of the following interactive components: the ordinance, trained enforcement personnel, sound measurement equipment, and prosecutorial mechanisms. If any of these components are not working in harmony, the NEP may not deliver the desired results. The NEP should be designed to reflect the realities of field enforcement, and should be tailored to the specific agency that will be conducting the enforcement, mindful of the differences between environmental compliance officers and police officers. This begins with the crafting of an ordinance that is appropriate for the unique conditions within the jurisdiction. The sound measurement equipment should be the least complicated necessary to take the specified measurements, and should be as durable as possible. Enforcement officers require proper training. Noncompliance must be addressed in a constructive and meaningful manner. A number of jurisdictions with active noise enforcement programs have been interviewed to determine the status of the program, and their strengths and weaknesses have been honestly assessed.

11:40

5aNS11. Noise impact evaluation method for supermarket sites. Gregory Tocci and Rose Mary Su (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA 01776, gtocci@cavtocci.com)

A large food supermarket chain is currently in a large store-building program involving several project managers who identify prospective sites, retain A/E design services, and obtain local permits for construction. There is a variety of environmental and planning issues that sometimes need special consideration; among these is environmental noise. Supermarket management has asked that consultants representing each discipline, and who work with project managers on specific projects as needed, issue guidelines as to when special consideration of their disciplines is required. This presentation describes a simple method developed for the supermarket owner that allows their project managers to judge whether there is a need for special consideration of environmental noise in store design by an acoustical consulting firm.

Session 5aPAa

Physical Acoustics: Wave Propagation: Homogeneous Media

Doug Meegan, Chair
Applied Research Laboratory, University of Austin, Austin, Texas 78713-8029

Contributed Papers

8:00

5aPAa1. Measurements and simulations of an asymmetric finite amplitude ultrasonic field. Kirk D. Wallace, Mark R. Holland, and James G. Miller (Dept. of Phys., Washington Univ. in St. Louis, 1 Brookings Dr., Campus Box 1105, St. Louis, MO 63130, kirk.wallace@wustl.edu)

Transducer arrays used routinely for medical imaging generate diffracting finite amplitude ultrasonic fields that are typically asymmetric and intricate. The goal of this work is to validate a Burgers equation enhanced, nonlinear angular spectrum simulation approach with experimental measurements of an asymmetric finite amplitude ultrasonic field. A one-dimensional transducer array with a 3:2 aspect ratio, separate azimuth and elevation foci, and a nonuniform aperture apodization was immersed in a watertank and was driven by a broadband pulse. A 0.6-mm-diam hydrophone receiver was mechanically scanned to obtain detailed maps of the harmonic content and spatial distribution of the finite amplitude ultrasonic field in nine transverse (00 to 160 mm, evenly spaced) and two orthogonal meridian plane cross sections. The fundamenal (2.3 MHz) component of the field, measured at face of the source, served as input for the numerical simulation (written in MATLAB and run on a notebook computer). Comparisons were performed at the fundamental, second, third, fourth, and fifth harmonics. Overall, excellent agreement was observed between experimental measurements of the nonlinear ultrasonic field generated with an asymmetric array transducer and the Burgers equation enhanced, nonlinear angular spectrum simulations. [Work supported by NIH R01 HL72761.]

8:15

5aPAa2. Numerical three-dimensional solution of broadband acoustic pulses through media with power-law attenuation. Margaret G. Wismer (Dept. of Elec. Eng., Bucknell Univ., Lewisburg, PA 17837, wismer@bucknell.edu)

Acoustic waves in tissues and weakly attenuative fluids often have an attenuation parameter, α, satisfying $\alpha=\alpha_0\omega^y$, in which ω is the applied frequency and y is between 1 and 2. This power-law attenuation is not predicted by the classical thermoviscous wave equation, and recent research has led to a number of modified viscous wave equations in which the third term usually consists of a convolution operator or a fractional spatial or temporal derivative. These wave equations are obtained by taking into account the requisite wave velocity dispersion predicted by the attenuation in order for the signals to be causal. In this paper, acoustic waves undergoing power-law attenuation are modeled by a slight modification to the thermoviscous wave equation, in which the time derivative of the viscous term is replaced by a fractional time derivative. This new equation satisfies the power-law formulation for lossy waves. An explicit time-domain, finite-element formulation leads to a stable algorithm capable of simulating three-dimensional broadband acoustic pulses propagating through attenuative and dispersive media. Results are given for pulse propagation through layered lossy media, and it is shown how attenuation affects the transmission and reflection of broadband signals.

8:30

5aPAa3. Numerical simulation of wave propagation in air including the effects of molecular relaxation and relative humidity. Mark S. Wochner, Anthony A. Atchley, and Victor W. Sparrow (Grad. Prog. in Acoust., Penn State Univ., University Park, PA 16802)

The research reported is directed at developing a computational model for assessing the importance of nonlinear propagation in the near field of jet engines. In order to attain accurate results, more complete models of acoustic wave propagation are required. The goal of the current research is to develop a two-dimensional time domain solution to the Navier–Stokes equations that includes the effects of classical absorption, relative humidity, and molecular relaxation of diatomic nitrogen and oxygen. Using fluid dynamics equations from Pierce [*Acoustics: An Introduction to Its Physical Principals and Applications* (ASA, New York, 1989)], a system of equations is created and solved using finite difference schemes. Benchmark cases will be discussed to show that the code is capable of reproducing the physical phenomena associated with atmospheric absorption and the relative importance of the various attenuation mechanisms will be discussed. [Work supported by ONR.]

8:45

5aPAa4. Numerical modeling of infrasound propagation at very long distance. Pierre-Franck Piserchia and Roger Roche (BP12, Bruyeres le chatel 91680, France, pierre-franck.piserchia@cea.fr)

Compliance with the CTBT in the atmosphere will be monitored by a world-wide network of infrasound stations consisting of 60 stations equipped with microbarographs in order to measure small changes in the air pressure in the frequency range 0.02 to 4 Hz. They are characterized by a good sensitivity, and by a large dynamic. By the application of array techniques, it is possible to determine the direction of pressure pulses caused by small explosions in the atmosphere, as well as shock waves caused by supersonic aircraft or meteorites. To take into account the nonlinear phenomena at the source and during the propagation, we are developing a numerical approach to solve the Euler nonlinear equation. In a first step, in the linear domain, this method is compared with two other numerical modeling approaches based on the ray tracing technique and the parabolic approach. In our test case, the source is on the ground and generates a 1-Pa pressure pulse centered at the frequency of 0.1 Hz. We considered an infrasound propagation over a distance of 500 km and an atmosphere height of 200 km. In a further step, the source level will be increased to study nonlinear phenomena.

9:00

5aPAa5. Terrain effects on acoustic pulse propagation. Donald G. Albert (USA ERDC-CRREL, 72 Lyme Rd., Hanover, NH 03755, Donald.G.Albert@erdc.usace.army.mil)

Outdoor sound propagation was measured at a variety of locations to determine terrain and environmental effects on acoustic pulses. Small explosive charges were detonated at a height of 1.5 m above the ground, and the resultant acoustic pulses were digitally recorded using pressure sensors located 10 to 100 m from the source. The experiments were conducted in an open area over concrete, grass, or snow; through a forest or thick tropical vegetation; and along a building-lined street in an artificial village. The measurements reveal that waveform changes occur at the shortest measurement ranges and increase as the propagation distance increases. These changes include amplitude attenuation, pulse width, and waveform shape and are caused by the differing ground conditions. The highest peak amplitudes at 100 m were recorded over concrete, and the lowest over snow-covered ground, with peak pressure decay rates of 24- and 32-dB/decade distance in these two situations. In addition, multiple scattering and reflected arrivals were visible in the tropical, forest, and urban areas. [Work funded by U.S. Army.]

9:15

5aPAa6. Modes and quasimodes for outdoor sound propagation at night. Roger Waxler (NCPA, Univ. of Mississippi, University, MS 38677, rwax@olemiss.edu)

Sound propagation outdoors at night is characterized by sound-speed profiles which are downward refracting in the first few hundred meters of the atmosphere but upward refracting at higher altitudes. The downward refraction causes sound to be ducted along the ground. Due to the upward refraction up high, however, this ducting is imperfect in that sound can leak away into the upper atmosphere. It has been shown that for most frequencies the sound is attenuated by the ground before it can leak out of the duct, effectively stabilizing the duct. For these frequencies a modal expansion has been developed which gives an efficient and physically transparent model for low-angle propagation. There are, however, certain resonant frequency bands in which the leaking can be significant. For these frequencies the modal expansion is not sufficient, but needs to be augmented by the addition of quasimodes. The quasimodes, while not strictly speaking modes, have the same form as the modes. A straightforward method for determining the quasimodes will be presented. The result is a modal expansion with uniform validity from frequencies of a few Hz up to a few kHz.

FRIDAY MORNING, 28 MAY 2004 ROYAL BALLROOM B, 9:45 A.M. TO 12:00 NOON

Session 5aPAb

Physical Acoustics: Wave Propagation: Time Reversal Methods

Evgenia A. Zabolotskaya, Chair

Department of Mechanical Engineering, University of Texas at Austin, 1 University Station, Austin, Texas 78712-0292

Contributed Papers

9:45

5aPAb1. Application of time reversal acoustics focusing for nonlinear imaging ms. Armen Sarvazyan and Alexander Sutin (Artann Labs., Lambertville, NJ, armen@artannlabs.com)

Time reversal acoustic (TRA) focusing of ultrasound appears to be an effective tool for nonlinear imaging in industrial and medical applications because of its ability to efficiently concentrate ultrasonic energy (close to diffraction limit) in heterogeneous media. In this study, we used two TRA systems to focus ultrasonic beams with different frequencies in coinciding focal points, thus causing the generation of ultrasonic waves with combination frequencies. Measurements of the intensity of these combination frequency waves provide information on the nonlinear parameter of medium in the focal region. Synchronized stirring of two TRA focused beams enables obtaining 3-D acoustic nonlinearity images of the object. Each of the TRA systems employed an aluminum resonator with piezotransducers glued to its facet. One of the free facets of each resonator was submerged into a water tank and served as a virtual phased array capable of ultrasound focusing and beam steering. To mimic a medium with spatially varying acoustical nonlinearity a simplest model such as a microbubble column in water was used. Microbubbles were generated by electrolysis of water using a needle electrode. An order of magnitude increase of the sum frequency component was observed when the ultrasound beams were focused in the area with bubbles.

10:00

5aPAb2. 3D shear wave generation in soft tissues using the time reversal kaleidoscopes. Delphine Palacio, Jeremy Bercoff, Gabriel Montaldo, Mickael Tanter, Mathias Fink (Laboratoire Ondes et Acoustique, 10 rue Vauquelin, 75005 Paris, France), Armen Sarvazyan, and Alexander Sutin (Artann Labs., Inc., Trenton, NJ)

The time reversal kaleidoscope provides elegant possibilities for both temporal and spatial concentration of acoustic energy in a wide frequency band. Using such ultrasound focused beams, it is possible to generate mechanical sources radiating low frequency shear waves inside tissues. Shear waves are of great interest to investigate tissue viscoelastic properties. In most techniques relying on this concept, generation and imaging of shear waves are realized with a single ultrasonic probe, limiting the capabilities concerning the generation of the shear source. Here we propose a system able to generate and image shear waves in a full 3D volume with any kind of polarization. The focused beam is generated by the time reversal kaleidoscope. By combining the concepts of time reversal mirrors and chaotic reverberating cavities, the time reversal kaleidoscope is able to focus ultrasound beams in a large 3D volume with only some tens of transducers and presents equivalent performances than conventional 2D matrices made of thousands of transducers. The kaleidoscope is triggered by an ultrafast imaging system able to image the resulting shear waves. *In vitro* results in tissue mimicking phantoms are presented. Chaotic cavities of different geometries are tested and compared. [Work supported in part by NIH grant.]

10:15

5aPAb3. Broadband time reversed acoustic focusing and steering system. Alexander Sutin (Artann Labs. and Stevens Inst. of Technol., NJ), Armen Sarvazyan (Artann Labs., Lambertville, NJ), Gabriel Montaldo, Delphine Palacio, Jeremy Bercoff, Mickael Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, 75005 Paris, France)

We present results of experimental testing and theoretical modeling of a time reversal acoustic (TRA) focusing system based on a multifaceted aluminum resonator with 15 piezoceramic transducers glued to the resonator facets. One of the facets of the resonator, a pentagon with characteristic dimension of about 30 mm, was submerged into a water tank and served as a virtual phased array which provided ultrasound focusing and beam steering in a wide frequency band (0.7–3 MHz). Ultrasonic pulses with different carrier frequencies and various complex waveforms were focused; the focal length was varied in the range of 10–55 mm and the focused beam was steered in a range of angles of ± 60 deg. The amplitude of the signal in the focal region reached 40 MPa. A theoretical model was based on an assumption that the radiating part of the resonator works as a phase conjugation screen for a spherical wave radiated from the focal point. Theoretical dependencies of the field structure on the position of the focus point and ultrasound frequency are in a good agreement with experimental results. TRA based focusing of ultrasound has numerous applications in medical diagnostics, surgery and therapy. [Work supported by NIH grant.]

10:30

5aPAb4. Acoustic tomographic array simulation. Sandra L. Collier, David A. Ligon, and John M. Noble (U.S. Army Res. Lab., 2800 Powder Mill Rd., Adelphi, MD 20783-1197)

Acoustic travel-time tomography of the atmosphere allows one to retrieve the virtual temperature and wind velocity fields, and to monitor their evolution in time. The temperature and wind velocity fields in a horizontal slice near the ground have been successfully retrieved by several research groups. One future direction for acoustic tomography of the atmosphere is the retrieval of detailed time-varying volumetric wind and temperature fields. The ability to retrieve these fields is strongly dependent upon the system configuration, e.g., the number and placement of sources and sensors, the geometry of the array, the environmental effects, and the type of the source. A sensitivity analysis for the retrieval of the temperature and wind velocity fields as a function of sensor configuration has been performed using an acoustic tomographic array simulator. We discuss the simulation model and results of the sensitivity analysis to some baseline propagation cases.

10:45

5aPAb5. Time reversal for source detection in urban environment. Lanbo Liu and Donald G. Albert (USA Cold Regions Res. and Eng. Lab., 72 Lyme Rd., Hanover, NH 03755-1290, lanbo.liu@erdc.usace.army.mil)

This paper applies the time reversal method to acoustic source detection in an outdoor urban environment. Experimental measurements were conducted in a full-scale artificial village to determine the effect of buildings on sound propagation outdoors. Explosive charges were detonated to produce acoustic pulses that were digitally recorded by sensors scattered throughout the village. We use the two-dimensional finite difference time domain (FDTD) method to compute synthetic time traces for acoustic propagation within the artificial village. The FDTD model predicted arrival times and amplitude levels that are in fair agreement with the measured data at most microphone measurement locations. Using the simulated time traces and the measured data from only nine stations and back propagating them into the model, the sound energy refocuses in the vicinity of the true source location. Our time reversal experiment confirms that the phase information is more critical than sound pressure levels and that using information acquired only at non-line-of-sight (NLOS) locations is sufficient to obtain accurate source locations. The results demonstrate that time reversal can potentially be used for fast source location in a complex urban terrain and noisy acoustic background, with without requiring LOS sensors. [Work funded by U.S. Army.]

11:00

5aPAb6. The time-reversal effect at sound scattering by a rough surface. Iosif Fuks and Konstantin Naugolnykh (Univ. of Colorado/Zeltech, Boulder, CO)

The scattering of the time-reversed wave by a rough surface is considered in the present paper. Using the Green function theory the point source field scattering by the rough surface in the Kirchhoff approximation is considered and the equation for the time-reversed and retransmitted back to the rough surface wave is obtained. The spatial distribution and the intensity spectrum of the scattered wave are evaluated for the Gaussian model of the rough surface and their features are examined. The modifications of the spectrum of the time-reversed wave as a function of the relation of the propagation time of the wave to the source-receiver array to the characteristic period of the surface roughness variation are discussed.

11:15

5aPAb7. Characterization of scattering object with the decomposition of the time-reversal operator: Theory and experiment. Jean-Gabriel Minonzio, Claire Prada (Laboratoire Ondes et Acoustique, CNRS, ESPCI, 10 rue Vauquelin, 75005 Paris France, claire.prada-julia@espci.fr), David H. Chambers (Lawrence Livermore Natl. Lab., Livermore, CA 94551), Dominique Clorennec, and Mathias Fink (CNRS, 75005 Paris, France)

The decomposition of the time-reversal operator provides information on the scattering medium. It has been shown [Chambers, J. Acoust. Soc. Am. **109**, (2001)] that a small spherical scatterer is in general associated with four eigenvalues and eigenvectors of the time-reversal operator. In this paper, the 2D problem of scattering by a cylinder measured by a linear array of transducers is considered. Experimental results are obtained for wires of different diameters and different materials. It is shown how the singular-value distribution and the singular vectors depend on the elastic wave speeds cL, cT, the density and the radius of each wire. These results offer a new perspective towards solution of the inverse problem by determining more than scattering contrast using conventional array processing like that used in medical ultrasonic imaging.

11:30

5aPAb8. Revisiting the Stokes relations in a time-reversal cavity: Suppression of intraplate echoes induced by a titanium plate. Jean-Francois Aubry, Francois Vignon, Michael Tanter, Gabriel Montaldo (Laboratoire Ondes et Acoustique, ESPCI, Université Paris VII, U.M.R. C.N.R.S. 7587, 10 rue Vauquelin, 75005 Paris, France, jf.aubry@espci.fr), and Mathias Fink (Université Paris VII, 75005 Paris, France)

When focusing through plates or tubes the presence of multiple interfaces induces reflected wavefronts that follow the main wavefront. Adaptive focusing techniques can be used to cancel the echoes. For that purpose, two linear arrays of transducers have been placed on each side of a titanium plate. Three propagation operators have been acquired: transmission from one array to the other, and two reflection operators acquired by each array. In this work, two adaptative focusing methods have been used to cancel the echoes: first, they have been suppressed with a time-reversal mirror, using the two arrays' cavity surrounding the plate. Second, the echoes have been canceled by using the inverse filter technique, inverting the transmission operator. Thus, the inverse filter achieves echo cancellation by using only the transmitted fields, whereas time reversal also requires the reflected fields. It is shown how transmission and reflection operators are related by the Stokes relations in a matrix formalism. These relations clearly exhibit how the inverse filter takes advantage of the reflections in the medium. An iterative mathematical resolution of these equations yields a new way to invert the transmission operator.

11:45

5aPAb9. Time reversal imaging of noise sources inside a reverberant room. Guillemette Ribay, Cédric Roux, Julien de Rosny, and Mathias Fink (Lab. Ondes et Acoustique, Université Denis Diderot, UMR CNRS 7587, ESPCI, 10 rue Vauquelin, 75005 Paris, France)

In a reverberant room, the multiple reflected echoes make imaging of acoustical random sources difficult. In this talk, we give an original solution based on Time Reversal Principles. Basically, in a first step, the acoustic field due to a noise source is recorded by a set of microphones. In a second step, all the recorded signals are time-reversed and re-emitted by loudspeakers that now replace the microphones. We observe a strong enhancement of the mean intensity around the initial source position. In a first part, we highlight the formal link between time-reversal of pulsed and noise sources in reverberant media. Especially we deduce that the signal to noise ratio depends only on the number of pairs of microphone/loud speaker for noise sources. Numerical simulations confirm these predictions. In a second part, we present experimental results performed in a $5*3*3=m$ and reverberant room. The intensity map of noise acoustic sources with a working bandpass of [100–3000] Hz has been performed. The same experiment has been carried out with two coherent or incoherent sources. Moreover, the frequency dependence of the resolution obtained with this technique is studied. Finally, we focus on the SNR with respect to the bandpass and central frequency.

FRIDAY MORNING, 28 MAY 2004 IMPERIAL BALLROOM B, 8:00 A.M. TO 12:00 NOON

Session 5aPPa

Psychological and Physiological Acoustics: Poster Session II

William P. Shofner, Chair

Parmly Hearing Institute, Loyola University of Chicago, 6525 North Sheridan Road, Chicago, Illinois 60626

Contributed Papers

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

5aPPa1. Spatial release from informational masking. Brad Rakerd (Dept. of Audiol. and Speech Sci., Michigan State Univ., East Lansing, MI 48824) and Neil L. Aaronson (Michigan State Univ., East Lansing, MI 48824)

A new method for investigating spatial release from informational masking was developed and employed in two experiments. The new method is computer controlled and efficient. It employs the versatile coordinate response measure speech stimulus set [Bolia *et al.*, J. Acoust. Soc. Am. **107**, 1065 (2000)]. The experiments were conducted in an anechoic room, with a primary loudspeaker in front of the listener and a secondary loudspeaker at 60 deg to the right. Target messages were presented from the primary speaker only. For a standard, distractor messages, simultaneous with the target, were also presented from the primary speaker only. Spatial release was measured by presenting the distractors from both primary and secondary speakers with a temporal offset. Experiment 1 fixed the offset (secondary leading, +4 ms) and varied the number of distractors (1 to 3) and the target-to-distractor ratio (−12 to +4 dB). Masking release, sometimes as large as 10 dB, was found for all combinations of these variables. Experiment 2 varied the offset over a wide range of values. Substantial release from masking was found for both positive and negative offsets, but only in the range in which speech echoes are suppressed (<50 ms). [Work supported by NIDCD grant DC 00181.]

5aPPa2. Localization of frequency-separated bands of noise in the median plane. Gerald Ng and H. Steven Colburn (Hearing Res. Ctr., Boston Univ., 44 Cummington St. #427, Boston, MA 02215, geraldng@bu.edu)

Six subjects reported the perceived location of two sets of noise stimuli presented through speakers in the median-sagittal plane. Two sets of stimuli were generated by passing flat (1 to 16 kHz) Gaussian noise through multiple, equal-amplitude bandpass filters. The control stimulus included all frequencies between 1 and 16 kHz. The first set contained two to six equally spaced, 1/3-octave-wide noise bands. The second set contained five equally spaced noise bands that varied in common bandwidth. Loudspeakers in the medial-sagittal plane were located at six frontal locations (−30, −15, 0, 15, 30, 45, with 0 directly in front) and two rear locations (120, 180). The experiments were conducted in a sound-treated room with the speakers concealed from view. Stimuli were 200 ms in duration and presented from one speaker per trial. Subjects used a pen to mark perceived stimulus locations on preprinted coordinate diagrams. Despite noticeable intersubject differences, most subjects localized stimuli with five or six 1/3-octave-wide bands as well as control stimuli. Many subjects showed steady decreases in accuracy as the bandwidths of five bands were reduced. HRTFs were recorded for each listener, and several model algorithms are being evaluated. [Work supported by NIDCD: Grants R01 DC00100 and P30 DC04663.]

5aPPa3. Localization of multiple-band noises in the median plane. Gerald Ng and H. Steven Colburn (Hearing Res. Ctr., Boston Univ., 44 Cummington St. #427, Boston, MA 02215, geraldng@bu.edu)

Six subjects reported the perceived location of two sets of noise stimuli presented through speakers in the median-sagittal plane. Two sets of stimuli were generated by passing flat (1 to 16 kHz) Gaussian noise through multiple, equal-amplitude bandpass filters. The control stimulus included all frequencies between 1 and 16 kHz. The first set contained two to six equally-spaced, $\frac{1}{3}$-octave-wide noise bands. The second set contained five equally spaced noise bands that varied in common bandwidth. Loudspeakers in the medial-sagittal plane were located at six frontal locations (−30, −15, 0, 15, 30, 45, with 0 directly in front) and two rear locations (120, 180). The experiments were conducted in a sound-treated room with the speakers concealed from view. Stimuli were 200 ms in duration and presented from one speaker per trial. Subjects used a pen to mark perceived stimulus locations on preprinted coordinate diagrams. De-

5a FRI. AM

spite noticeable intersubject differences, most subjects localized stimuli with five or six $\frac{1}{3}$-octave-wide bands as well as control stimuli. Many subjects showed steady decreases in accuracy as the bandwidths of five bands were reduced. HRTFs were recorded for each listener, and several model algorithms are being evaluated. [Work supported by NIDCD Grants R01 DC00100 and P30 DC04663.]

5aPPa4. Improved BTE hearing-aid directivity using a directional microphone array. Douglas L. Jones (Dept. of Elec. and Computer Eng., Univ. of Illinois at Urbana–Champaign, 1406 W. Green St., Urbana, IL 61801, dl-jones@uiuc.edu), Michael E. Lockwood (Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801), Charissa R. Lansing (Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801), and Albert S. Feng (Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801)

Extraction of speech in noise is of great importance to hearing-impaired listeners. Directional microphones are incorporated in some hearing aids to improve noise rejection through increased directivity. A symmetric cardioid response is created either with a single directional microphone, or using beamforming with two omnidirectional microphones. The head-related transfer function (HRTF), however, introduces an asymmetry that cannot be exploited by a linear array of omnidirectional microphones. A new BTE array consisting of a gradient directional microphone with nulls in the front–back vertical plane and two omnidirectional microphones exploits the asymmetry of the HRTF to obtain almost 2 dB better directivity than the best cardioid. HRTFs measured on KEMAR with this array were transformed to the frequency domain, where directivity-maximizing coefficients in each band were derived. The Articulation-Index (AI) weighted directivity gain of this optimal three-element directional array was 6.4 dB greater than a single omnidirectional microphone on the BTE, whereas the directivity gain of the HRTF-optimized two-omni beamformer was 4.6 dB, and the optimal free-field cardioid placed on the head yielded 4.4 dB. [Work partially supported by NIH (NIDCD) under Grant No. 1 R01 DC005762-01A1.]

5aPPa5. Effect of source location and listener location on ILD cues in a reverberant room. Antje Ihlefeld and Barbara G. Shinn-Cunningham (Hearing Res. Ctr., Boston Univ., 677 Beacon St., Boston, MA 02215, ihlefeld@bu.edu)

Short-term interaural level differences (ILDs) were analyzed for simulations of the signals that would reach a listener in a reverberant room. White noise was convolved with manikin head-related impulse responses measured in a classroom to simulate different locations of the source relative to the manikin and different manikin positions in the room. The ILDs of the signals were computed within each third-octave band over a relatively short time window to investigate how reliably ILD cues encode source laterality. Overall, the mean of the ILD magnitude increases with lateral angle and decreases with distance, as expected. Increasing reverberation decreases the mean ILD magnitude and increases the variance of the short-term ILD, so that the spatial information carried by ILD cues is degraded by reverberation. These results suggest that the mean ILD is not a reliable cue for determining source laterality in a reverberant room. However, by taking into account both the mean and variance, the distribution of high-frequency short-term ILDs provides some spatial information. This analysis suggests that, in order to use ILDs to judge source direction in reverberant space, listeners must accumulate information about how the short-term ILD varies over time. [Work supported by NIDCD and AFOSR.]

5aPPa6. Itakura distance measure used to study speech intelligibility in a room using white noise. Jorge Sommerhoff, Jorge Cardenas, Victor Poblete, and Jose Luis Barros (Instituto de Acustica, Universidad Austral de Chile, Campus Miraflores, Valdivia, Chile, jsommerh@uach.cl)

The recognition process of isolated words use distance measures to determine if the word corresponds to one of its vocabulary. In this investigation Itakura distance measure was used to compare the segments of white noise of two signals with different relationships of direct and reverberant energy densities. Interesting data is obtained from the distance measure in the function of the position of the microphone in the room in which a white noise source is continuously working. These results could be useful to develop a new methodology to measure the word intelligibility in a room. There were carried out some comparative measurements with Rasti and intelligibility test. [Work supported by DID–UACh.]

5aPPa7. Perceptual scaling of room reverberation. Pavel Zahorik (Dept. of Psychol. and Brain Sci., Univ. of Louisville, Louisville, KY 40292, pavel.zahorik@louisville.edu)

Recent evidence suggests that reverberant energy can provide listeners with important spatial information regarding the distance of a sound source. However, relatively little is known about the perceptual attributes of the reverberation itself, and how these attributes may be related to physical properties of the environment that also potentially impact perceived spatial location. Here, perceived similarity among 15 reverberant rooms simulated using virtual auditory space techniques was examined. Room size and surface absorption properties were varied, along with aspects of the virtual simulation including the use of individualized head-related transfer function (HRTF) measurements and properties of the room acoustic simulation. Seven listeners rated perceived similarity on a 100-point scale between all possible pairs of simulated rooms using a speech source signal. Multidimensional scaling techniques were used to estimate scales of perceived room reverberation. Although the resulting scales were complex and somewhat unique to individual listeners, it is clear that the perceptual effects of manipulating properties of the reverberant sound are much larger than the effects due to either nonindividualized HRTFs or nonoptimal room simulation methods. [Work supported by NIDCD.]

5aPPa8. Disrupting the build-up of the precedence effect. Richard L. Freyman (Dept. of Commun. Disord., Univ. of Massachusetts, Amherst, MA 01003, RLF@comdis.umass.edu) and Rachel Keen (Univ. of Massachusetts, Amherst, MA)

When pairs of clicks are presented with a brief delay to one loudspeaker, the perception of a single fused image builds up as the clicks are repeated. If lead and lag loudspeakers are then reversed, the precedence effect resets, i.e., two auditory images are often heard immediately after the switch. However, Djelani and Blauert [Acta Acust. (Beijing) **87**, 253–261 (2001)] found that a brief reversal in lag location did not reset the built-up fusion at the original spatial configuration. The current experiment determined whether the build-up effect would be disrupted or slowed by interleaving variations in lead–lag stimuli within a train of repeated clicks. During the standard train the lead was presented from 45 deg left and the lag from 45 deg right, the click pairs were presented at a rate of 4/s, the number of clicks ranged from 3 to 9, and the delay between lead and lag varied from 2 to 14 ms across trials. Results show that the variation that disrupted the build-up most consistently was when clicks from only the lead loudspeaker were interleaved. This condition simulated the absence of expected reflections. [Work supported by NIDCD DC01625.]

5aPPa9. A real-time virtual auditory system for spatially dynamic perception research. Jacob W. Scarpaci and H. Steven Colburn (Hearing Res. Ctr. and Dept. of Biomed Eng., Boston Univ., 44 Cummington St., Boston, MA 02215, scarpaci@bu.edu)

A Real Time Virtual Auditory System (RT-VAS) is being developed to provide a high-performance, cost-effective, flexible system that can dynamically update filter coefficients in hard real time on a PC. An InterSense head tracker is incorporated to provide low-latency head tracking to allow studies with head motion. Processing is done using a real-time Linux Kernel (RTAI kernel patch) which allows for precise processor scheduling, resulting in negligible time jitter of output samples. Output is calculated at a sample rate of 44.1 kHz and displayed using a National Instruments DAQ. Object oriented approach to system development allows for customizable input, position, and calculation routines as well as multiple independent auditory objects. Input and position may be calculated in real-time or read from a file. Calculation of output may include filtering with spatially sampled HRTFs or analytic models and head movements may be recorded to file. Limitations of the system are tied to the speed of the processor, thus complexity of experiments scales with speed of computer hardware. The current system handles multiple moving sources while tracking head position. Preliminary psychoacoustic results with head motion will be shown, as well as a demonstration of the system. [Work supported by NIH DC00100.]

5aPP10. Temporal integration of interaural-time differences carried by trains of 500-Hz tone bursts. Michael A. Akeroyd (MRC Inst. of Hearing Res., Glasgow Royal Infirmary, 16 Alexandra Parade, Glasgow G31 2ER, UK, maa@ihr.gla.ac.uk)

The just-noticeable difference for the interaural time difference (ITD) of a sound generally reduces as its duration is increased [e.g., T. Houtgast and R. Plomp, J. Acoust. Soc. Am. **44**, 807–812 (1968); E. Hafter and R. H. Dye, J. Acoust. Soc. Am. **73**, 644–651 (1983)]. The rate of this temporal integration is less than would be expected from a simple multiple-looks application of signal detection theory, as though one part of the sound is given greater weighting. The present experiment measured the contributions of the start, middle, or end of a sound to the detectability of its ITD. The stimulus was a 32-pip train of 500-Hz, 10-ms pips, diotic apart from some (2–16) target pips which carried the ITD to be detected and which were placed either at the start, middle, or end of the train. Integration rates were determined from psychometric functions measured using four normal-hearing listeners. The results showed that the rate was least for targets placed at the start or end but larger for targets at the middle. The data can be described using a weighted multiple-looks approach, based on an integration function of approximately 100-ms time constant together with an emphasis of the start, and end, of the sound.

5aPPa11. A shot-noise phase-opponency model for monaural detection of pure tones in noise at low frequencies. Yan Gai and Laurel H. Carney (Inst. for Sensory Res. & Dept. of Bioengineering & Neurosci., Syracuse Univ., Syracuse, NY 13244, mountfall@yahoo.com)

Phase-locked responses of low-frequency auditory-nerve (AN) fibers contain sufficient information for detection of a tone in noise. The phase-opponency (PO) model [Carney *et al.*, Acustica **88**, 334–347 (2002)] uses relative phase across fibers with different characteristic frequencies (CFs). The model's discharge rate is reduced when a tone is added to noise due to changes in the coincidence of the rate functions across AN CFs. In the current study, a shot-noise model for coincidence detection was used to implement the PO model. AN discharge times were simulated using an AN model with a discharge generator; therefore, internal noise was naturally included in the AN input fibers. Model results were compared to human data for detection of a 500-Hz tone in reproducible noises, which provided external noise. The aim was to test the hypothesis that the shot-noise PO model explains human performance in terms of thresholds, hit- and false-alarm rates, and to understand the performance for different reproducible noises. The ratio of the internal to external noise in the shot-noise PO model depended on combining responses of several identical but independent cells. [Work supported by NIDCD R01-01641.]

5aPPa12. Comparing spectral resolvability in chinchillas and human listeners using phase discrimination. William P. Shofner, Kathryn Sparks, Yuanxing Esther Wu, and Ellen Pham (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, wshofne@luc.edu)

A tone complex made of harmonic components that are added in cosine-starting phase can be discriminated from complexes comprised of identical harmonics that are added with random-starting phases. Phase discrimination occurs when unresolved harmonics interact within a single auditory channel. When harmonics are resolved, there is less interaction among components resulting in poorer phase discrimination performance. Thus, phase discrimination indirectly reflects spectral resolvability. Performance in a phase discrimination task was measured in chinchillas and human listeners to compare spectral resolvability between the two groups. Subjects discriminated a cosine-phase tone complex from random-phase tone complexes in a go/no-go behavioral paradigm. Tone complexes were comprised of a 250-Hz fundamental frequency and N consecutive higher harmonics, where N was 5, 10, 20, and 40. Performance was evaluated in terms of d'. The results show that the measured d' increased as N increased, and values of d' for each N condition were similar between chinchillas and human listeners. Values of the criterion for each N condition were also similar between chinchillas and humans. The results do not support the hypothesis that spectral resolvability is poorer in chinchillas, but suggest that resolvability is similar between the two groups. [Work supported by NIH/NIDCD.]

5aPPa13. Mechanisms of forward masking. Magdalena Wojtczak and Neal F. Viemeister (Dept. of Psych., Univ. of Minnesota, 75 East River Rd., Minneapolis, MN 55455, wojtc001@umn.edu)

Experiments involving forward masking were run with the aim of elucidating the mechanism underlying this phenomenon. Two possible mechanisms were considered: adaptation following masker offset, and persistence of excitation. In one experiment, detection of a 10-ms noise probe following a 150-ms noise masker was measured as a function of the level of a 20-ms burst of noise temporally centered on the probe. Model predictions were generated assuming either a reduction of gain applied to the probe (adaptation) or additivity of persisting masker excitation and excitation produced by the probe. Both models accounted for the data equally well. In the second experiment, listeners performed an image-centering task involving binaural presentation of a 4-kHz 10-ms probe. In one ear, the probe was preceded by a 150-ms noise forward masker. The level of the probe in the unmasked ear was adjusted until the image of the probe was centered. In general, the image was centered when the probe level in the unmasked ear was much lower (by 10–20 dB) than that in the masked ear. This result could be explained in terms of adaptation underlying forward masking but not in terms of persistence of excitation. [Work supported by NIH/NIDCD DC00683.]

5aPPa14. Discrimination of common-envelope signals: Comparison of static and dynamic frequency sounds. Lawrence Feth, Dean Hudson (Dept. of Speech and Hearing Sci., Ohio State Univ., Columbus, OH 43210, feth.1@osu.edu), Daniel Hack, Vivek Rajendran, and Ashok Krishnamurthy (Ohio State Univ., Columbus, OH 43210)

Listeners can match the rate of frequency change in an FM glide to the transition rate heard in a virtual frequency (VF) glide. A VF glide is perceived in a two-tone complex when the amplitude of one tone increases, while that of the other tone decreases, over the duration of the signal. This phenomenon is robust for signals spanning five or more auditory filter bands (ERBs). Static two-tone complexes that approximate the end points of the VF glide are discriminable for components separated by more than 20 Hz, but less than five ERB. Preliminary work on discrimination of VF glides versus FM glides processed to have the same envelope as the VF signal led to estimates of just discriminable bandwidth less than one ERB (Rajendran *et al.*, 2003). Center frequencies ranged from 500 to 4000 Hz, but duration was fixed at 250 ms. The preliminary study has been extended to include a range of signal durations from 25 to 250 ms, and to determine the maximum resolution bandwidth for each combination of duration and center frequency. Predictions of the perceptual spectral centroid model will be compared to the results from well-practiced listeners.

5aPPa15. Scaling the perceived fluctuation strength of frequency-modulated tones. Florian Wickelmaier and Wolfgang Ellermeier (Dept. of Acoust., Aalborg Univ., Fredrik Bajers Vej 7 B5, 9220 Aalborg East, Denmark, fw@acoustics.dk)

Fluctuation strength is one of the major psychoacoustic variables considered in sound-quality evaluation. Zwicker and Fastl [*Psychoacoustics* (Springer, Berlin, 1999)] summarize recommendations for its computation, which have already been implemented in various software applications, even though the data basis is rather limited. In particular, the dependency of fluctuation strength on modulation frequency and modulation depth has seemingly never been tested in a factorial design. Therefore, in experiment I both of these factors were varied simultaneously in order to create 54 different frequency-modulated sinusoids. The task of the subjects was to directly estimate the perceived magnitude of fluctuation strength. The results do not conform well with the prevalent model of fluctuation strength. In experiment II this finding was further investigated by varying only one factor at a time. The results show that large individual differences, particularly in the effect of modulation frequency, persist. Thus, in experiment III the interaction of both factors was analyzed on an individual basis. By employing a 2AFC procedure, matches in fluctuation strength were obtained. The results suggest that most listeners are not able to integrate modulation frequency and modulation depth additively into a unidimensional percept. [Work supported by Centercontract on Sound Quality, Aalborg University.]

5aPPa16. Effect of noise on the detection of intensity increments. Walt Jesteadt, Lance Nizami, Stephen T. Neely, and Kim S. Schairer (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131, jesteadt@boystown.org)

The detection of an increment in a longer duration tone is often used as a measure of intensity resolution. Many studies of increment detection have used background noise, but few have varied the level of the longer duration tone relative to the noise or have compared results obtained with or without the noise. To explore this space in more detail, thresholds were obtained by adaptive tracking from four subjects for detection of a 50-ms increment in a 350-ms, 4000-Hz tone presented in quiet or in broadband noise of 0, 10 and 20 dB spectrum level. Tone levels ranged from 15 to 75 dB SPL in quiet and from approximately 0 to 30 dB above the tone's detection threshold in the noise. Increment detection thresholds expressed as values of the difference in overall level of the tone resulting from addition of the increment increase as tone sensation levels decrease and are greater in noise than in quiet, but increase only slightly as a function of noise level. The data are well fitted by a model that assumes a specific form of nonlinearity motivated by loudness-matching data. [Work supported by NIDCD.]

5aPPa17. Loudness recalibration at short ISI: A closer look. Yoav Arieh (Dept. of Psych., Montclair State Univ., Montclair, NJ 07043, ariehy@mail.montclair.edu), Jennifer R. Mailloux (Mary Washington College, Fredericksburg, VA 22401-5358), and Lawrence E. Marks (John B. Pierce Lab., New Haven, CT 06519)

The loudness of a moderate level tone is substantially reduced when preceded by a louder tone of the same frequency (loudness recalibration). The reduction in loudness depends on the interstimulus interval (ISI) between the tones. Arieh and Marks (2003) reported that loudness starts to decline only when ISI exceeds about 200 m. One possible explanation is that loudness reduction takes 200 ms to develop. Alternatively, the delay in loudness reduction could reflect the combination of two processes: short-term facilitation and longer-term suppression. To test this hypothesis, subjects compared the loudness of a roving-level 500-Hz tone to a 60-dB 2500-Hz target tone presented, first, 100 ms after an 80-dB 2500-Hz recalibration tone and, subsequently, with the recalibration tone omitted; by omitting the recent recalibration tone, we hoped to observe the long-term effects on loudness of earlier recalibration tones. As expected, the loudness of the target was unaffected when it followed the recalibration tone by 100 ms, but loudness declined significantly when the recalibration tone was subsequently omitted. These results are consistent with the hypothesis that at short ISI, a recalibration tone sets off long-term suppression of loudness that is initially offset by short-term facilitation.

5aPPa18. Simple reaction time to narrow-band and broadband noise. Eva Wagner, Mary Florentine (Inst. of Hearing, Speech & Lang. and SLPA Dept. (106A FR), Northeastern Univ., 360 Huntington Ave, Boston, MA 02115, e.wagner@neu.edu), Søren Buus, and Joseph McCormack (Northeastern Univ., Boston, MA 02115)

The present study evaluates the relationship between loudness matches and simple reaction times (RT) in six normal-hearing subjects. Loudness matches between a narrow-band noise (125 Hz wide) and a broadband noise (1500 Hz) centered at 1000 Hz were made at levels from near threshold to near 100 dB SPL. For the same noises and level range, RT were also measured. In agreement with previous loudness-matching studies, as SPL increased the level difference between the noises needed to maintain equal loudness first increased, to around 10 dB at moderate SPLs, and then decreased. Except for one listener, the RT data show the same pattern. The level difference needed to maintain RT the same to the two noises first increased and then decreased. These results show that RT is closely related to loudness, but not to sensation level. If RT depended on sensation level, the level difference between the two noises needed to achieve equal RT would not change with SPL because the difference in sensation level between two sounds is constant as a function of SPL. Overall, the present data provide strong support for the contention that simple RT depends strongly on loudness. [Supported by NIH/NIDCD Grant No. R01 DC 02241.]

5aPPa19. Relative intelligibility of dynamically extracted transient versus steady-state components of speech. J. R. Boston, Sungyub Yoo, C. C. Li, Amro El-Jaroudi (Dept. of Elec. Eng., Univ. of Pittsburgh, Pittsburgh, PA 15261, boston@engr.pitt.edu), J. D. Durrant, Kristie Kovacyk, and Stacey Karn (Univ. of Pittsburgh, Pittsburgh, PA 15261)

Consonants are recognized to dominate higher frequencies of the speech spectrum and to carry more information than vowels, but both demonstrate quasi-steady state and transient components, such as vowel to consonant transitions. Fixed filters somewhat separate these effects, but probably not optimally, given diverse words, speakers, and situations. To enhance the transient characteristics of speech, this study used time-

varying adaptive filters [Rao and Kumaresan, IEEE Trans. Speech Audio Process. **8**, 240–254 (2000)], following high-pass filtering at 700 Hz (well-known to have minimal effect on intelligibility), to extract predominantly steady-state components of speech material (CVC words, NU-6). The transient component was the difference between the sum of the filter outputs and the original signal. Psychometric functions were determined in five subjects with and without background noise and fitted by ogives. The transient components averaged filtered speech energy, but PBmax was not significantly different (nonparametric ANOVA) from that of either the original or highpass filtered speech. The steady-state components yielded significantly lower PBmax (p 3D 0.003) despite their much greater energy, as expected. These results suggest a potential approach to dynamic enhancement of speech intelligibility. [Work supported by ONR.]

5aPPa20. Contribution of consonant versus vowel information to sentence intelligibility by normal and hearing-impaired listeners. T. Zachary Burkle, Diane Kewley-Port, Larry Humes, and Jae Hee Lee (Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, tzachbur@indiana.edu)

The purpose of this study was to examine the contribution of information provided by vowels versus consonants to sentence intelligibility in young normal-hearing (YNH) and elderly hearing-impaired (EHI) listeners. Sentences were presented in three conditions, with either the vowels or the consonants replaced with speech shaped noise, or unaltered. Sentences from male and female talkers in the TIMIT database were selected. EHI subjects listened at 95 dB SPL, and YNH subjects at both 95 and 70 dB SPL. Subjects listened to each sentence twice and were asked to repeat the entire sentence after each presentation. Words were scored correct if identified exactly. Average performance for unaltered sentences was greater than 94%. Vowel-present conditions were always significantly more intelligible than consonant-present conditions, similar to data reported by Cole and colleagues [Proceedings of ICASSP, 1996]. Across groups, performance in the vowel-present conditions exceeded that in the consonant-present conditions by 14% to 40%, although EHI subjects performed more poorly than YNH subjects. In contrast to written English, vowels in spoken language carry more information about sentences than consonants for both normal and hearing-impaired listeners. [Work supported by NIDCD-02229.]

5aPPa21. Speech reception thresholds in various interference conditions. Suzanne P. Carr and H. Steven Colburn (Hearing Res. Ctr. and Dept. of Biomed Eng., Boston Univ., Boston, MA 02215)

Speech intelligibility is integral to human verbal communication; however, our understanding of the effects of competing noise, room reverberation, and frequency range restriction is incomplete. Using virtual stimuli, the dependence of intelligibility threshold levels on the extent of room reverberation, the relative locations of speech target and masking noise, and the available frequency content of the speech and the masking noise is explored. Speech-shaped masking noise and target sentences have three spectral conditions: wideband, high pass above 2-kHz, and low pass below 2-kHz. The 2-kHz cutoff was chosen to approximately bisect the range of frequencies most important in speech, and the high pass noise condition simulates high-frequency hearing loss. Reverberation conditions include a pseudo-anechoic case, a moderately reverberant "classroom" case, and a very reverberant "bathroom" case. Both binaural and monaural intelligibility are measured. Preliminary results show that source separation decreases thresholds, reverberation increases thresholds, and low frequency noise reverberates more in the rooms, contributing to increasing thresholds along with the effects of the upward spread of masking. The energetic effects of reverberation are explored. [Work supported by NIH DC00100.]

5aPPa22. Revised audibility-noise variance of speech recognition sensitivity model. Hannes Müsch (Sound ID, 3430 W. Bayshore Rd., Palo Alto, CA 94303, hmuesch@soundid.com) and Søren Buus (Northeastern Univ., Boston, MA 02115)

The speech recognition sensitivity (SRS) model [H. Müsch and S. Buus, J. Acoust. Soc. Am. **109**, 2896–2909 (2001)] was used to predict normal-hearing listeners' consonant-recognition performance in a nonsense-syllable context as measured in nine independent studies. By attempting to fit these data, the relation between the variance of the SRS model's audibility noise σ_N^2 and the speech-excitation to noise-excitation ratio SNR_E in the auditory periphery was optimized. The best predictions were obtained when the audibility-noise variance was directly proportional to the noise-excitation power. Based on this finding it is suggested that the relation between audibility-noise variance and SNR_E that was proposed in the original SRS paper be replaced by $10\log_{10}(\sigma_{N_i}^2) = -SNR_{E_i} + B + 10\log_{10}((z_j - z_i)/(19\mathrm{Barks}))$, where $\sigma_{N_i}^2$ is the audibility-noise variance in the band reaching from critical-band rate z_i to z_j, SNR_{E_i} is the speech-excitation to noise-excitation ratio in that band expressed in dB, and B is a fitting constant that will be in the order of 13 dB. The model provided good predictions of the data with model parameters that are consistent with those of earlier studies. [Work supported by NIH/NIDCD Grant R01DC00187.]

5aPPa23. Lateral suppression preserves speech intelligibility at high intensities. James A. Bashford, Jr., Richard M. Warren, and Peter W. Lenz (Dept. of Psych., Univ. of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201, bashford@uwm.edu)

The intelligibility of narrowband speech decreased 23% when its level was raised from 45 to 75 dB. However, when flanking bands of low-pass and high-pass white noise were added, intelligibility at the higher speech level recovered by as much as 65%. Recovery appears to be due to lateral suppression counteracting overloading effects of auditory-nerve (AN) firing-rate saturation at high speech intensities. Findings supporting this hypothesis include: (1) the absence of intelligibility enhancement at the lower speech level; (2) a greater effect of the higher frequency flanking noise band at a low noise level; and (3) equivalent effects of continuous versus gated flanking noise, inconsistent with firing-rate adaptation. In addition, there was no intelligibility increase when the flanking noise and narrow-band speech were delivered to opposite ears of listeners. These behavioral results are consistent with previous physiological observations and models of lateral suppressive interactions occurring in the lower auditory pathway. It appears that the noise-induced intelligibility recovery is produced via lateral inhibition of saturated AN-fiber input to neurons of the cochlear nucleus, with a possible additional contribution from mechanical (two-tone) suppression evoked within the cochlea and likely not involving olivocochlear feedback. [Work supported by NIH.]

5aPPa24. Development of a topic-related sentence corpus for speech perception research. Karen S. Helfer and Richard L. Freyman (Univ. of Massachusetts, Amherst, MA 01003)

A large sentence corpus has been developed for use in speech recognition research. Sentences ($n=881$, three scoring words per sentence) were developed under 23 topics. In the first phase of development subjects rated each individual scoring word for relatedness to its given topic on a Likert scale. Next, two groups of young, normal-hearing listeners ($n=16$/group) listened and responded to the recordings of the sentences (spoken by a female talker) presented with one of two types of maskers: steady-state noise ($S{:}N=-13$ dB) or two other females speaking random sentences ($S{:}N=-8$ dB). Each subject responded to half of the sentences with topic supplied and half with no topic supplied. Data analyses focused on addressing two questions: whether supplementation of topic would be more important in the presence of the speech masker versus the noise masker, and how the degree of relatedness of each key word to the topic influenced the effect of topic on recognition. The data showed little dif-

ference in how beneficial the topic was for speech versus noise maskers. Moreover, there was a complex relationship between effect of topic, type of masker, and position of the word in the sentence. [Work supported by NIDCD DC01625.]

5aPPa25. Immersive simulation of hearing loss and auditory prostheses. Patrick M. Zurek and Joseph G. Desloge (Sensimetrics Corp., 48 Grove St., Somerville, MA 02144, pat@sens.com)

Simulation of hearing loss is useful for demonstrating the communication challenges facing hearing-impaired people. However, current simulations, most of which are only recordings, do not actually elevate thresholds; i.e., they do not simulate hearing loss, *per se*. The hearing loss simulator described in this talk is immersive; the user's detection thresholds for ambient sounds are shifted by a prescribed degree. This threshold shift is achieved through a combination of passive attenuation (from muff-type hearing protectors) and additive masking noise (introduced by within-muff earphones). Acoustic signals picked up by microphones near each ear are processed through bandpass AGC channels and delivered via the earphones to complete the simulation of frequency-dependent hearing loss and loudness recruitment. Preliminary results validating the accuracy of specified threshold shift will be presented, along with speech-reception data comparing simulated with actual hearing losses. Subjective reactions of users engaged in one-on-one conversation suggest that strong feelings of communication disability are engendered by even moderate degrees of simulated hearing loss. The system, which is capable of simulating any degree of recruiting hearing loss along with hearing aids or cochlear implants, can provide effective interactive demonstrations of both auditory communication handicap and rehabilitation options. [Work supported by NIDCD.]

FRIDAY MORNING, 28 MAY 2004 — CONFERENCE ROOM E, 8:45 TO 11:15 A.M.

Session 5aPPb

Psychological and Physiological Acoustics: More Complex Stimuli

Elizabeth A. Strickland, Cochair
Audiology and Speech Sciences, Purdue University, 1353 Heavilon Hall, West Lafayette, Indiana 47907

Jungmee Lee, Cochair
Speech and Hearing Sciences, City University of New York Graduate Center, 365 Fifth Avenue, New York, New York 10016

Contributed Papers

8:45

5aPPb1. Uncertainty and confusion in temporal masking. C. Formby (Div. of Otolaryngol.-HNS, Univ. of Maryland School of Medicine, 16 S. Eutaw St., Ste. 500, Baltimore, MD 21201, cformby@smail.umaryland.edu) and T. Zhang (Univ. of Maryland, College Park, MD 20742)

In a landmark study, Wright *et al.* [Nature **387**, 176–178 (1997)] reported an apparent backward-masking deficit in language-impaired children. Subsequently, these controversial results have been influential in guiding treatments for childhood language problems. In this study we revisited Wright *et al.*'s temporal-masking paradigm to evaluate listener uncertainty effects. Masked detection was measured for 20-ms sinusoids (480, 1000, or 1680 Hz) presented at temporal positions before, during, or after a gated narrowband ($W=600$–1400 Hz) masker. Listener uncertainty was investigated by cueing various stimulus temporal properties with a 6000-Hz sinusoid presented either ipsi- or contra-lateral to the test ear or bilaterally. The primary cueing effect was measured in the backward-masking condition for a contralateral cue gated simultaneously with the on-frequency 1000-Hz signal. The resulting cued masked-detection threshold was reduced to quiet threshold. No significant cueing effects were obtained for other signal temporal positions in the masker nor for any off-frequency signal conditions. These results indicate that (1) uncertainty can be reduced or eliminated for on-frequency backward masking by cueing the signal and (2) the deficit reported by Wright *et al.* for language-impaired children may reflect uncertainty and confusion rather than a temporal-processing deficit *per se*. [Research supported by NIDCD.]

9:00

5aPPb2. Effect of masker onset asynchrony on overshoot in simultaneous masking. Andrzej Miskiewicz (Dept. of Speech-Lang. Pathol. and Audiol. (106A FR), Inst. of Hearing, Speech and Lang., Northeastern Univ., Boston, MA 02115), Soren Buus, and Mary Florentine (Northeastern Univ., Boston, MA 02115)

This study examines how overshoot is influenced by asynchrony between the onsets of an on-frequency center-band noise masker (CB) and a flanking noise band (fringe). Thresholds were measured for 2-ms tones at 5 kHz gated on 2 ms after the onset of the CB. Measurements were made with only the CB (4590–5464 Hz), and with either a low fringe (1900–4590 Hz) or a high fringe (5500–11 000 Hz) added to the CB. The fringe came on between 500 ms before and 100 ms after the CB onset. Because previous studies demonstrated that overshoot is produced only by masker components outside the auditory filter centered at the signal frequency, overshoot is taken as the increase in threshold caused by adding the fringe. Results show that considerable overshoot, 7–11 dB, is produced by a high fringe gated on within 10 ms before or after CB onset; overshoot is maximum when the fringe and CB come on simultaneously. Maximum overshoot with a low fringe is less than 4 dB. The finding that a high fringe increases threshold even when it comes on shortly after the signals end suggests that overshoot is governed primarily by central mechanisms. [Work supported by NIH/NIDCD Grant R01DC00187.]

9:15

5aPPb3. Audibility of time-varying signals in time-varying backgrounds: Model and data. Brian C. J. Moore and Brian R. Glasberg (Dept. of Exp. Psych., Univ. of Cambridge, Downing St., Cambridge CB2 3EB, England, bcjm@cus.cam.ac.uk)

We have described a model for calculating the partial loudness of a steady signal in the presence of a steady background sound [Moore *et al.*, J. Audio Eng. Soc. **45**, 224–240 (1997)]. We have also described a model for calculating the loudness of time-varying signals [B. R. Glasberg and B. C. J. Moore, J. Audio Eng. Soc. **50**, 331–342 (2002)]. These two models have been combined to allow calculation of the partial loudness of a time-varying signal in the presence of a time-varying background. To evaluate the model, psychometric functions for the detection of a variety of time-varying signals (e.g., telephone ring tones) have been measured in a variety of background sounds sampled from everyday listening situations, using a two-alternative forced-choice task. The different signals and backgrounds were interleaved, to create stimulus uncertainty, as would occur in everyday life. The data are used to relate the detectability index, d', to the calculated partial loudness. In this way, the model can be used to predict the detectability of any signal, based on its calculated partial loudness. [Work supported by MRC (UK) and by Nokia.]

9:30

5aPPb4. Sinusoidal spectral modulation masking period patterns. Aniket Saoji and David A. Eddins (Psychoacoustic Lab., Ctr. for Hearing and Deafness, Dept. of Commun. Disord. and Sci., Univ. at Buffalo, Buffalo, NY)

Knowledge of the internal auditory representation of spectral envelope features is critical to the understanding of auditory perception by listeners with normal and impaired hearing. Previous investigators have used forward and simultaneous masking to derive masking patterns that reveal the influence of lateral suppression and other factors on the internal representation of various spectral envelopes (e.g., narrowband noise and steady-state vowels). The present study followed a general approach, in which the internal representation of spectral shape was estimated on the basis of spectral masking patterns obtained with sinusoidal spectral modulation. Simultaneous- and forward-masked thresholds were measured for probe signals as a function of spectral modulation frequency (cycles/octave), phase, and depth (peak-valley contrast in dB) superimposed on a noise carrier (400–6400 Hz). For relatively low modulation frequencies (<1 cycle/octave), masking patterns closely mimicked features of the external spectrum. For very high modulation frequencies (>2 cycles/octave), masking patterns revealed reduced spectral contrast, consistent with the limited frequency selectivity of the auditory system. For intermediate modulation frequencies (1–2 cycles/octave), considerable spectral sharpening was evident. The three-dimensional space defined by spectral modulation frequency, depth, and internal contrast will be discussed in terms of predicting the internal representation of arbitrary complex spectral envelopes.

9:45

5aPPb5. The effect of modulation maskers on the detection of second-order amplitude modulation in the absence and presence of notched noise. Rosalie M. Uchanski (Dept. of Otolaryngol., Washington Univ. School of Medicine, St. Louis, MO 63110), Brian C. J. Moore, and Brian R. Glasberg (Univ. of Cambridge)

Second-order amplitude modulation (AM) can be detected despite the absence of a component at the second-order modulation frequency in the modulation spectrum. This may depend on the detection of a modulation component at the second-order frequency produced by nonlinearities in the auditory system, or on the detection of a "beat" at the output of the modulation filter centered at the first-order rate. To assess these possibilities, and to explore the role of off-frequency listening, first- and second-order AM detection thresholds were determined for six listeners with normal hearing, in the absence and presence of a notched noise centered at the carrier frequency of 4 kHz. First-order thresholds were measured for $f1 = 2$, 10, 16, and 50 Hz. Second-order thresholds were measured for $f1 = 16$ and $f2 = 2$ Hz, and for $f1 = 50$ and $f2 = 10$ Hz. Then, using a second-order modulation depth giving about 80%-correct detection, performance for detecting second-order AM was measured in the presence of a modulation masker centered at either the first- or second-order rate. In the absence of notched noise, both modulation maskers impaired performance, for both second-order rates. In the presence of notched noise, for most listeners there was little effect of either modulation masker. [Work supported by MRC.]

10:00

5aPPb6. Auditory discrimination of binary sequences. Stanley Sheft and William A. Yost (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, ssheft@luc.edu)

Auditory sequence discrimination was investigated for stimuli generated with binary modulation coding schemes. Both amplitude- and frequency-shift keying (ASK and FSK) were used to convert 16-event binary patterns into a series of pulses. Discrimination was measured as a function of the number of binary events that differed between the standard and comparison sequences with temporal locations of the deviant pulses randomly selected on each trial. Across conditions, pulse duration ranged from 4 to 64 ms. In the ASK scheme, the carrier was wideband noise with off-key amplitudes scaled by a factor ranging from 0.0 to 0.5. Performance improved with increasing either pulse duration, amplitude shift, or the number of deviant pulses from one to eight. These trends are consistent with envelope processing by a modulation filterbank which discards phase information above low modulation rates. In the FSK scheme, binary patterns were coded by pulse-frequency differences centered at 1.25 kHz. As with ASK, FSK results showed performance improving with pulse duration and number of deviant pulses. FSK findings suggest that the low-pass results obtained with ASK envelope coding may reflect a general characteristic of sequence discrimination that is not restricted solely by auditory envelope processing. [Work supported by NIDCD.]

10:15

5aPPb7. Perceptually related analysis of time-frequency patterns via a hearing model (Sottek), a pattern-measurement algorithm ("relative approach") and a window-deconvolution algorithm. Wade Bray (HEAD acoustics, Inc., 6964 Kensington Rd., Brighton, MI 48116)

The presence of patterns (time structure and/or tonal structure) in sound situations elicits a different analytic process by the human hearing system than when sound situations are without pattern. In the presence of pattern, two issues can interfere with the effectiveness of conventional measurement methods: the human hearing is not acting as an absolute magnitude measurer so conventional measurements may not, or only poorly, resolve the important criteria; and pattern information that is subjectively highly significant is frequently of low magnitude amid general high-level structure that is not a direct part of the subjective pattern-response problem yet carries significance. Several automotive sound quality and information technology time-data recordings involving various patterns, of different strengths, were analyzed using conventional, psychoacoustic, and specialized pattern-sensitive techniques. In general, the advanced perception-related methods proved superior in quantifying "patterned" sound situations according to subjective impression.

10:30

5aPPb8. Touch-induced scraping sounds and texture perception. M. Ercan Altinsoy (Inst. of Commun. Acoust., IGSN, Ruhr Univ. Bochum, 44780 Bochum, Germany, ercanaltinsoy@rub.de)

Sound is often the result of human–object interaction (touch, striking, scraping, etc.) and conveys to the listener required information about physical attributes of the interaction, and spatial properties of the sound event, e.g., location, geometry. People are highly skilled in using touch-produced sounds to identify texture properties. Loudness and pitch of the touch-induced scraping sounds are the important psychoacoustical determinants of the texture perception [S. J. Lederman, Perception **8**, 93–103 (1979)]. In this study, psychophysical experiments were conducted to investigate the relationship between auditory properties (loudness and pitch) and texture perception. In order to control the psychoacoustical parameters, the scraping sounds were synthesized in computer environment. The results of the experiments show that subjects were able to differentiate the roughness of textures on the basis of pitch and loudness differences. Decreasing pitch tends to result in increasing magnitude of perceived roughness.

10:45

5aPPb9. Revealing criteria which underlie judgments of auditory unpleasantness. Karin Zimmer, Wolfgang Ellermeier, and Christian Schmid (Dept. of Acoust., Aalborg, Univ., Fredrik Bajers Vej 7 B-5, DK-9440 Aalborg Denmark, kaz@acoustics.dk)

The usefulness of probabilistic choice models in attaining metric scales of sensation magnitude is demonstrated by scaling the perceived unpleasantness of environmental sounds. Seventy-four subjects made pairwise comparisons of the unpleasantness of 12 binaurally recorded sounds. The stimuli varied considerably in loudness, timbre, and temporal envelopes. A preference-tree choice model accounted well for the structure underlying the data, indicating that subjects changed criteria when evaluating different sound pairs, and that these criterion changes combined in a lawful way, making it possible to measure unpleasantness on a ratio scale across the entire set of sounds investigated. The sounds could be grouped according to their (nonacoustical) intrusiveness, and loudness. A subsequent multiple-regression analysis showed that in the subgroups identified, a combination of psychoacoustical sharpness and roughness explained the unpleasantness-scale values very well; $r^2_{\mathrm{corr}}=0.96$. It is thus illustrated that choice models provide an explicitly stated theory of the observers' decision behavior, with built-in checks of the consistency of judgments, and statistical tests to validate if the attempt at (ratio) scale construction succeeded. This is in contrast to techniques of "direct" scaling, e.g., magnitude estimation, in which the scale level has to be assumed *bona fide*. [Work supported by Centercontract on Sound Quality, Aalborg University.]

11:00

5aPPb10. Coherence and the Speech Intelligibility Index. James Kates (GN ReSound, Boulder Res. Group, 3215 Marine St., Rm. W161, Boulder, CO 80309, jkates@gnresound.dk) and Kathryn Arehart (Univ. of Colorado, Boulder, CO 80309)

Noise and distortion reduce the sound quality in hearing aids, but there is no established procedure for calculating sound quality in these devices. This presentation introduces a new intelligibility and sound-quality calculation procedure based on the Speech Intelligibility Index [ANSI S3.5-1997]. The SII involves measuring the signal-to-noise ratio (SNR) in separate frequency bands, modifying the estimated noise levels to include auditory masking, and computing a weighted sum across frequency of the modified SNR values. In the new procedure, the estimated signal and noise levels are replaced with estimates based on the coherence between the input and output signals of the system under test. Coherence is unaffected by linear transformations of the input signal, but is reduced by nonlinear effects such as additive noise and distortion; the SII calculation is therefore modified to include nonlinear distortion as well as additive noise. For additive noise, the coherence calculation gives SII scores identical to those computed using the standard procedure. Experiments with normal-hearing listeners using additive noise, peak-clipping distortion, and center-clipping distortion are then used to relate the computed coherence SII scores with the subjects' intelligibility and quality ratings. [Work supported by GN ReSound (JMK) and the Whitaker Foundation (KHA).]

FRIDAY MORNING, 28 MAY 2004 IMPERIAL BALLROOM B, 8:00 A.M. TO 12:00 NOON

Session 5aSC

Speech Communication: Poster Session IV

Lori L. Holt, Chair

Department of Psychology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213

Contributed Papers

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

5aSC1. The discrimination and the production of English vowels by bilingual Spanish/English speakers. Sandra Levey (Dept. of Speech-Lang.-Hearing Sci., Lehman College of the City Univ. of New York, 250 Bedford Park Blvd. W., Bronx, NY 10468)

The discrimination of English vowels in real and novel words by 40 bilingual Spanish/English participants was examined. Their discrimination was compared with that of 40 native monolingual English participants. Participants were 23–36 years of age (mean 25.3; median 25.0). Stimuli were presented within triads in an ABX paradigm. This categorial discrimination paradigm was selected to avoid labeling, allowing participants to indicate categories to which stimuli belonged. Bilingual participants' productions of vowels in real words used in the discrimination task were judged by two independent listeners. The goal was to determine the degree of correlation between discrimination and production. Vowels were studied as these segments present second language learners with more difficulty than consonants. Discrimination difficulty was significantly greater for bilingual participants than for native English participants for vowel contrasts and novel words. Significant errors also appeared in the bilingual participants' production of certain vowels. English vowels absent from Spanish presented the greatest difficulty, while vowels similar to those in Spanish presented the least difficulty. Earlier age of acquisition, absence of communication problems, and greater percentage of time devoted to communication in English contributed to greater accuracy in discrimination and production. [Work supported by PSC-CUNY.]

5aSC2. Proficient bilinguals require more information for vowel identification than monolinguals. Alexandra S. Lopez and Catherine L. Rogers (Dept. of Commun. Sci. & Disord., Univ. of South Florida, 4202 E. Fowler Ave., PCD1017, Tampa, FL 33620-8100)

Even proficient bilinguals have been shown to experience more difficulty understanding speech in noise than monolinguals. One potential explanation is that bilinguals require more information than monolinguals for phoneme identification. We tested this hypothesis by presenting gated, silent-center vowels to two groups of listeners: (1) monolingual American English speakers and (2) proficient Spanish–English bilinguals, who spoke unaccented or mildly accented English. To create the stimuli, two American English speakers were recorded as they read the following items: ''beeb, bibb, babe, bebb, babb,'' and ''bob.'' Duration-preserved silent-center versions of three tokens of each item were created by retaining varying amounts of the CV and VC transitions (10, 20, 30, or 40 ms) and attenuating the remainder of the vowel center to silence. Duration-neutral versions of silent-center tokens were created by lengthening or shortening the silent portion to match the tokens vowel duration to the average for all the tokens. Listeners identified the unedited (full vowel), duration-preserved, and duration neutral silent-center tokens in a six-alternative forced-choice task. The two groups of listeners identified the unedited tokens with similar accuracy. In the silent-center conditions, however, the bilinguals identified the stimuli less accurately than the monolinguals. [Work supported in part by NIH-NIDCD Grant No. 1R03DC005561-01A1.]

5aSC3. Discrimination of synthesized English vowels by American and Korean listeners. Byunggon Yang (English Dept., Dongeui Univ., 24 Kayadong, Pusanjingu, Pusan 614-714, Republic of Korea)

This study explored the discrimination of synthesized English vowel pairs by 27 American and Korean, male and female listeners. The average formant values of nine monophthongs produced by ten American English male speakers were employed to synthesize the vowels. Then, subjects were instructed explicitly to respond to AX discrimination tasks in which the standard vowel was followed by another one with the increment or decrement of the original formant values. The highest and lowest formant values of the same vowel quality were collected and compared to examine patterns of vowel discrimination. Results showed that the American and Korean groups discriminated the vowel pairs almost identically and their center formant frequency values of the high and low boundary fell almost exactly on those of the standards. In addition, the acceptable range of the same vowel quality was similar among the language and gender groups. The acceptable thresholds of each vowel formed an oval to maintain perceptual contrast from adjacent vowels. Pedagogical implications of those findings are discussed.

5aSC4. Acoustic cues in the perception of second language speech sounds. Anna A. Bogacka (School of English, Adam Mickiewicz Univ., al. Niepodleglosci 4, 61-874 Poznan, Poland)

The experiment examined to what acoustic cues Polish learners of English pay attention when distinguishing between English high vowels. Predictions concerned the influence of Polish vowel system (no duration differences and only one vowel in the high back vowel region), salience of duration cues and L1 orthography. Thirty-seven Polish subjects and a control group of English native speakers identified stimuli from heed-hid and who'd-hood continua varying in spectral and duration steps. Identification scores by spectral and duration steps, and $F1/F2$ plots of identifications, were given as well as fundamental frequency variation comments. English subjects strongly relied on spectral cues (typical categorical perception) and almost did not react to temporal cues. Polish subjects relied strongly on temporal cues for both continua, but showed a reversed pattern of identification of who'd-hood contrast. Their reliance on spectral cues was weak and had a reversed pattern for heed-hid contrast. The results were interpreted with reference to the speech learning model [Flege (1995)], perceptual assimilation model [Best (1995)] and ontogeny phylogeny model [Major (2001)].

5aSC5. Vowel normalization for accent: An investigation of perceptual plasticity in young adults. Bronwen G. Evans and Paul Iverson (Dept. of Phonet. and Linguist., Univ. College London, Wolfson House, 4, Stephenson Way, London NW1 2HE, UK, bron@phon.ucl.ac.uk)

Previous work has emphasized the role of early experience in the ability to accurately perceive and produce foreign or foreign-accented speech. This study examines how listeners at a much later stage in language development—early adulthood—adapt to a non-native accent within the same language. A longitudinal study investigated whether listeners who had had no previous experience of living in multidialectal environments adapted their speech perception and production when attending university. Participants were tested before beginning university and then again 3 months later. An acoustic analysis of production was carried out and perceptual tests were used to investigate changes in word intelligibility and vowel categorization. Preliminary results suggest that listeners are able to adjust their phonetic representations and that these patterns of adjustment are linked to the changes in production that speakers typically make due to sociolinguistic factors when living in multidialectal environments.

5aSC6. Exploring the intelligibilty of foreign-accented English vowels when ''English'' is ill-defined. Rikke Louise Bundgaard-Nielsen and Ocke-Schwen Bohn (English Dept., Aarhus Univ., Aarhus, Denmark)

Many studies of foreign-accented speech have been conducted in *second* language settings in which learners are assumed to be exposed to a relatively homogeneous non-native sound system. However, *foreign* language learners, who learn an additional language in a setting where this language is not the primary medium of communication, are frequently exposed to a range of varieties of the target language which may differ considerably with respect to their sound systems. The present study examined and compared the intelligibility of English monophthongs produced by two speaker groups: Native Danes who had learned English as a foreign language (with exposure to different native and non-native varieties) and native English speakers from Australia, the US, and the UK. Ten native Canadian-English listeners, who were familiar with native and non-native accents of English, identified the 11 monophthongs of English produced by the speaker groups in a /bVt/ context. As expected, the listeners' error patterns were specific for each speaker group. However, reduced intelligibility was observed for much the same vowels irrespective of speaker group. Our results suggest that one source of problems in learning the sounds of English is the heterogeneity of English vowel systems in addition to transfer from the native language.

5aSC7. ERP indices of speech processing in bilinguals. Karen Garrido, Miwako Hisagi, and Valerie Shafer (Speech and Hearing Sci., City Univ. of New York–Grad. Ctr., 365 Fifth Ave., New York, NY 10016)

The specific aim of this project was to examine perceptual differentiation of the vowel contrast [I-E] by early versus late bilingual Spanish–English speakers, compared to monolingual English speakers. Neither of these phonetic segments occur as phonemes in Spanish while they constitute phonemic contrast in English. For Spanish listeners, these vowels may be heard as variants of the Spanish phonemes [i] and [e], respectively. The formation of perceptual categories may be affected by exposure to two languages in which the phonemic status of phonetic segments differs. In the current study, electrophysiological measures (mismatched negativity—MMN) and behavioral measures (AX discrimination and identification) were employed to examine perception of a nine-step continuum resynthesized from natural tokens. Preliminary results indicate that early bilinguals show somewhat more categorical perception than monolinguals (i.e.,

sharper identification boundaries) of the nine-step continuum. MMN results, obtained during a task in which subjects were not attending to the auditory input, show clear MMNs for monolinguals for both cross- and within-category pairs. To date, results for early bilinguals show less clear indication of preattentive discrimination. The role of language experience and the relationship between the behavioral performance and the electrophysiological measures will be discussed.

5aSC8. Structural equation modeling applied to the prediction of word identification accuracy. Sumiko Takayanagi, Lynne E. Bernstein, and Edward T. Auer, Jr. (House Ear Inst., Los Angeles, CA 90057)

Five structural equation models (SEMs) were developed to examine explicitly the relative contribution of experiential (subjective frequency and word age of acquisition) and word similarity (lexical equivalence class size and phoneme equivalence class size) factors on word identification accuracy (WIA). WIA (percent word correct scores and word cost scores) was measured for 184 monosyllabic words under five presentation conditions: (1) visual-only speech; (2) high intelligibility vocoded auditory-only speech; (3) low intelligibility vocoded auditory-only speech; (4) high intelligibility vocoded audiovisual speech; and (5) low intelligibility vocoded audiovisual speech. The results showed that each factor can be treated as an isolated factor and can be measured explicitly. Furthermore, the relative strengths of their contributions varied as a function of intelligibility and the estimation power of the predictor variables. In addition, two models were developed to estimate audiovisual (conditions 4 and 5) WIA from experiential, visual- and auditory-similarity factors. Only the SEM of condition 4 fit the data, and the experiential and auditory factors contributed equally, but the visual factor did not contribute much to the WIA in this model. Advantages and limitations of SEM will be discussed.

5aSC9. Perceptual assimilation of French and German vowels by American English monolinguals: Acoustic similarity does not predict perceptual similarity. Winifred Strange, Erika Levy, and Robert Lehnholf, Jr. (Speech and Hearing Sci., City Univ. of New York—Grad. Ctr., 365 Fifth Ave., New York, NY 10016)

Previous research in our laboratory has demonstrated that the perceived similarity of vowels across languages is not always predictable from the closeness of their target formant values in $F1/F2/F3$ space. In this study, perceptual similarity was established using a task in which 11 American English (AE) monolinguals were presented multiple tokens of 9 French vowels and 14 North German vowels (in separate blocks) produced in citation-form /hVb(a)/ (bi)syllables by native speakers. They selected 1 of 11 AE vowel responses to which each non-native vowel token was most similar, and rated its goodness on a 9-point Likert scale. Of special interest was the perceptual assimilation of front rounded French [y, oe] and German [y, Y, o/, oe] vowels. Acoustically, all six French and German vowels are more similar to front unrounded AE vowels. However, all six vowels were perceived to be more similar to back rounded AE vowels (range across vowels = 55% to 100%), although relatively poor exemplars. There were differences across languages in how the same vowel was assimilated (e.g., French /y/ assimilated to front AE vowels 13%, German /y/, 0%; French [oe] 3%, German [oe] 45%). There were also large individual differences in listeners assimilation patterns. [Work supported by NIDCD.]

5aSC10. Segmental differences in the visual contribution to speech intelligibility. Kuniko Nielsen (Dept. of Linguist., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095-1543, kuniko@humnet.ucla.edu)

It is well known that the presence of visual cues increases the intelligibility of a speech signal (Sumby and Pollack, 1954). Although much is known about segmental differences in visual-only perception, little is known about the contribution of visual cues to auditory–visual perception for individual segments. The purpose of this study was to examine (1) whether segments differ in their visual contribution to speech intelligibility, and (2) whether the contribution of visual cues is always to increase speech intelligibility. One talker produced triples of real words containing 15 different English consonants. Forced-choice word-identification experiments were carried out with these recordings under auditory–visual (AV) and auditory-only (A) conditions with varying S/N ratios, and identification accuracy for the 15 consonants was compared for A versus AV conditions. As expected, there were significant differences in the visual contribution for the different consonants, with visual cues greatly improving speech intelligibility for most segments. Among them, labio-dentals and interdentals show the largest improvement. Although individual perceivers differed in their performance, the results also suggest that for some consonants, the presence of visual cues can reduce intelligibility. In particular, the intelligibility of [r] decreased significantly in the AV condition, being perceived as [w] in most cases.

5aSC11. The effects of age on identification of temporally altered visual-only speech signals. Brent Spehar, Nancy Tye-Murray (Dept. of Otolaryngol., Central Inst. for the Deaf, Washington Univ. School of Medicine, Box 8115, St. Louis, MO 63130), and Mitchell S. Sommers (Washington Univ., St. Louis, MO 63130, msommers@artsci.wustl.edu)

The purpose of the present study was to investigate the effects of age on the ability to identify temporally altered visual speech signals (i.e., visual-only presentations). Young (18–26) and older (over age 65) adults identified words placed in final position within a carrier phrase. Digital signal processing was used to compress (speed) and expand (slow) both the carrier phrase and the target words by 33%. Thus, the resultant stimuli were either 67% or 133% of the original duration. Speaking rate (expanded, unmodified, and compressed) was manipulated within subjects. Overall performance for both older and younger adults was significantly reduced for the speeded, compared with either the unaltered or the slowed condition. Furthermore, although older adults had significantly lower speechreading scores than younger adults, they did not exhibit disproportionate declines for the temporally altered conditions. These results stand in contrast to previous findings with auditory-only presentations in which older adults show significantly greater declines than younger adults as a consequence of increased speaking rate. Taken together, these findings argue against a generalized slowing of information processing in older adults and instead point to modality-specific changes in temporal processing abilities.

5aSC12. The effects of auditory–visual vowel and consonant training on speechreading performance. Carolyn Richie and Diane Kewley-Port (Dept. of Speech & Hearing Sci., Indiana Univ., 200 S. Jordan Ave., Bloomington, IN 47405, carodavi@indiana.edu)

Recent work examined the effects of a novel approach to speechreading training using vowels, for normal-hearing listeners tested in masking noise [C. Richie and D. Kewley-Port, J. Acoust. Soc. Am. **114**, 2337 (2003)]. That study showed significant improvements in sentence-level speechreading abilities for trained listeners compared to untrained listeners. The purpose of the present study was to determine the effects of combining vowel training with consonant training on speechreading abilities. Normal-hearing adults were tested in auditory–visual conditions in noise designed to simulate a mild-to-moderate sloping sensorineural hearing loss. One group of listeners received training on consonants in monosyllable context, and another group received training on both consonants and vowels in monosyllable context. A control group was tested but did not receive any training. All listeners performed speechreading pre- and post-tests, on words and sentences. Results are discussed in terms of differences between groups, dependent upon which type of training was administered; vowel training, consonant training, or vowel and consonant training combined. Comparison is made between these and other speechreading training methods. Finally, the potential benefit of these vowel- and consonant-based speechreading training methods for rehabilitation of hearing-impaired listeners is discussed. [Work supported by NIHDCD02229.]

5aSC13. Visual contributions to talker attunement. Lawrence Brancazio (Dept. of Psych., Southern Connecticut State Univ., 501 Crescent St., New Haven, CT 06515 and Haskins Labs, New Haven, CT, brancazioL1@southernct.edu)

Listeners can attune to talker-specific speech patterns, as demonstrated by findings of greater accuracy in spoken-words-in-noise identification for words spoken by familiar voices [e.g., Nygaard and Pisoni, Percept. Psychophys. **60**, 355–376 (1998)]. In previous demonstrations of this effect, talker familiarity was produced by auditory speech exposure. The purpose of the present study was to examine the contribution of visible speech information to talker attunement. First, the auditory attunement effect was replicated: Participants were given extensive exposure to auditory words spoken by different talkers, and were subsequently tested on novel words in noise spoken by one of the familiar talkers and by a talker not previously heard. Identification of the words was more accurate for the familiar talker. This procedure was then extended to include video presentations of the talker's articulating face along with each word during the exposure phase. To control for familiarity with the faces, a separate condition included presentation of static images of the talkers' faces. Visual articulations were not presented during the identification-in-noise task. Thus, the experiment tests whether visual exposure to talker's articulatory patterns contributes to talker attunement, and, specifically, whether any such visual contributions to talker attunement generalize to auditory speech perception. [Work supported by NICHD.]

5aSC14. The influence of task on gaze during audiovisual speech perception. Julie Buchan, Martin Paré, Micheal Yurick, and Kevin Munhall (Queen's Univ., Kingston, ON, Canada, buchanj@psyc.queensu.ca)

In natural conversation, visual and auditory information about speech not only provide linguistic information but also provide information about the identity and the emotional state of the speaker. Thus, listeners must process a wide range of information in parallel to understand the full meaning in a message. In this series of studies, we examined how different types of visual information conveyed by a speaker's face are processed by measuring the gaze patterns exhibited by subjects watching audiovisual recordings of spoken sentences. In three experiments, subjects were asked to judge the emotion and the identity of the speaker, and to report the words that they heard under different auditory conditions. As in previous studies, eye and mouth regions dominated the distribution of the gaze fixations. It was hypothesized that the eyes would attract more fixations for more social judgment tasks, rather than tasks which rely more on verbal comprehension. Our results support this hypothesis. In addition, the location of gaze on the face did not influence the accuracy of the perception of speech in noise.

5aSC15. Cues to gender in children's speech. Suzanne Curtin and Scott Kiesling (Dept. of Linguist., Univ. of Pittsburgh, Pittsburgh, PA 15260, scurtin@pitt.edu)

Awareness of one's own gender emerges around 3 years and awareness that gender stays stable throughout life is evident by 4 years (Bee, 1998). This suggests that by 4 years of age, noticeable gender differences may emerge along a number of dimensions. The hypothesis tested here is that adults are able to identify the gender of 4-year-olds by voice quality alone. Sixteen four-year-olds were recorded saying the alphabet. Small portions of each recording were excised, and played to 40 adults. Adults were asked to identify the gender of the speaker. Subjects were able to correctly identify the gender of the child more often than chance. However, in the cases where a child's gender was incorrectly identified, pitch did not play a significant role. Rather, it appears that formant structure is the best predictor (although not perfect), of how a child's gender will be judged by voice (Perry, 2001). These results have important implications for our understanding of the linguistic cues that listeners use to identify the gender of speakers; they must be relying on phonetic cues that are much more subtle than gross pitch, lexical, phonological, or syntactic differences which are the usual provisions of language and gender research.

5aSC16. Role of a carrier sentence in native speakers' identification of Japanese vowel length. Yukari Hirata (Dept. of EALL, 13 Oak Dr., Hamilton, NY 13346, yhirata@mail.colgate.edu) and Stephen G. Lambacher (Univ. of Aizu, Aizu-Wakamatsu, Fukushima 965-8580, Japan)

This study examined the role of word-external contexts in native speakers' identification of phonemic vowel length in Japanese disyllabic nonwords. Three groups of native Japanese listeners ($n=20$, 20, and 19) were assigned to three conditions differing in word-external contexts. In intact condition, three types of target disyllables /mVmV/, /mV:mV/, and /mVmV:/ were spoken in a carrier sentence at slow and fast speaking rates. In the other conditions, the target disyllables of the intact condition were excised from the original carrier sentence (excised), and embedded in the carrier sentence of the other rate (mismatch). The accuracy for identifying the length of the two vowels was significantly higher for the intact (98.1%) than the excised (69.5%) and the mismatch condition (52.4%). This indicates that the presence of the carrier sentence with an appropriate rate was essential for accurate identification of vowels, consistent with Johnson and Strange [J. Acoust. Soc. Am. **72**, 1761–1770 (1982)]. Further analysis indicated that the perceptual errors made in the excised and the mismatch conditions were closely associated with the absolute duration overlap between the short vowels spoken at the slow rate and the long vowels spoken at the fast rate.

5aSC17. On the just-noticeable difference for tempo in speech. Hugo Quené (Utrecht Inst. of Linguist., Utrecht Univ., The Netherlands, hugo.quene@let.uu.nl)

The rate or tempo of speech modulates listeners' phonetic expectations, e.g., about VOT or about segment durations. In addition, variations in tempo contribute directly to speech communication, e.g., by expressing the communicative importance of speech fragments. It is not clear, however, how large tempo differences in speech have to be in order to be noticeable and to contribute to speech perception. This paper attempts to determine the jnd for tempo in speech, by means of detection of gradual changes in tempo. Participants listened to 10-s spoken texts, in which overall tempo was linearly accelerated (to 80%) or decelerated (to 120%) over a 5-s interval. They were instructed to press a button as soon as they detected a tempo change in the text; jnd's were computed from their response times. Preliminary results suggest that the average jnd's for tempo in speech are about 14% (acceleration) and 15% (deceleration). These values are far larger than those reported for tempo in music (about 3% to 6%), which suggests that only large changes in tempo are relevant for speech communication.

5aSC18. Individual differences in speech and nonspeech perception of frequency and duration. Matthew J. Makashay[a)] (Dept. of Linguist., Ohio State Univ., 222 Oxley Hall, 1712 Neil Ave., Columbus, OH 43210, matthew.makashay@na.amedd.army.mil)

This study investigates whether there are systematic individual differences in the perceptual weighting of frequency and durational speech cues for vowels and fricatives (and their nonspeech analogs) among a dialectally homogeneous group of speakers. Listeners performed AX discrimination for four separate types of stimuli: sine wave vowels, narrow-band fricatives, synthetic vowels, and synthetic fricatives. Duration and $F1$ frequency were manipulated for the vowels in *heed* and *hid*, and duration and frequency of the fricative centroid in the $F5$ region were manipulated for the fricatives in *bath* and *bass*. Dialect production and perception tasks were included to ensure that subjects were not from dissimilar dialects. Multidimensional scaling results indicated that there are subgroups within a dialect that attend to frequency and duration differently, and that not all listeners use these cues consistently across dissimilar phones. If subgroups can have different perceptions of speech despite similar productions, this questions the requirements for classifying dialect continua. Furthermore, the ratios of these subgroups changing over time can explain some language mergers and shifts. [a]Currently at Army Audiol. & Speech Ctr., Walter Reed Army Medical Ctr., Washington, D.C.

5aSC19. Acoustic–phonetic convergence among interacting talkers. Jennifer S. Pardo (Psych. Dept., Columbia Univ., 1190 Amsterdam Ave. Mail Code 5501, New York, NY 10027, jsp2003@columbia.edu)

Among other sources, acoustic–phonetic variability derives from both personal attributes of a talker and the social structure of a conversational setting. This research begins to detail the effects of such sources by examining the influence of conversational interaction on speech production. A set of talkers provided samples of speech before, during, and after participating in a paired conversational task. Using an A×B paradigm with a separate set of listeners, initial research found that paired talkers became more similar in phonetic repertoire, that this convergence persisted over a short delay, and that the degree of convergence was influenced by both the sex and task role of a talker. This paper presents detailed acoustic and talker scaling analyses from this conversational corpus in an attempt to chart the effects of personal and situational factors on speech variability. [Research supported in part by NIMH Grant No. 5F32MH64995.]

5aSC20. Effects of speaking rate on the acceptability of change in segmental duration within a phrase. Makiko Muto,[a] Hiroaki Kato (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan, makiko.muto@ruri.waseda.jp), Minoru Tsuzaki (ATR Spoken Lang. Translation Res. Labs., Kyoto 619-0288, Japan), and Yoshinori Sagisaka (Waseda Univ., Tokyo 169-0051, Japan)

To contribute to the naturalness criteria of speech synthesis, acceptability of changes in segment duration has been investigated. Previous studies showed context dependency of the acceptability evaluation such as *intraphrase positional effect*, where listeners were more sensitive to the phrase-initial segment duration than the phrase-final one. Such contextual effects were independent of the original durations of the segments tested [Kato *et al.*, J. Acoust. Soc. Am. **104**, 540–549 (1998)]. However, past studies used only normal-speed speech and temporal variation was limited. The current study, therefore, examined the contextual effect with a wide variety of speaking rates. The materials were three-mora phrases with either rising or falling accent that were spoken at three rates (fast, normal, and slow) with or without a carrier sentence. The duration of each vowel was either lengthened or shortened (10–50 ms) and listeners evaluated the acceptability of these changes. The results showed a clear speaking-rate effect in parallel with the intraphrase positional effect: the acceptability declined more rapidly as the speaking rate became faster. These results, along with those of Kato *et al.*, suggest that acceptability is evaluated based on the speaking rate rather than on the original duration itself. [Work supported by TAO, Japan.] [a]Currently at GITI, Waseda University.

5aSC21. Perceptual learning of talker-specific vowel space. Jennifer S. Queen (Dept. of Psych., Rollins College, Winter Park, FL 32789, jqueen@rollins.edu) and Lynne C. Nygaard (Emory Univ., Atlanta, GA 30322)

Previous research suggests that as listeners become familiar with a speaker's vocal style, they are better able to understand that speaker. This study investigated one possible mechanism by which this talker familiarity benefit arises. Listeners' vowel spaces were measured using a perceptual discrimination test both before and after they were trained to identify a group of speakers by name. Listeners identified either the same speakers whose vowels they discriminated or a different group of speakers. Differences in the learnability and the intelligibility of the two speaking groups were observed. The speaker group that was harder to identify also had vowels that were harder to discriminate. Changes in the listeners' vowel spaces were determined by examining multidimensional scaling solutions of their responses during the discrimination tests. All listeners became better at discriminating vowels. However, only listeners who heard different speakers during identification training and vowel discrimination exhibited a shift in their vowel spaces after training. This suggests that encountering new voices in an unrelated, nonlinguistic task acts to alter the perceptual context and may affect the structure of linguistic representation. Together, these results suggest a link between linguistic and nonlinguistic information in representations for spoken language.

5aSC22. Talker-specific auditory imagery during reading. Lynne C. Nygaard, Jessica Duke, Kathleen Kawar (Dept. of Psych., Emory Univ., Atlanta, GA 30322, lnygaar@emory.edu), and Jennifer S. Queen (Rollins College, Winter Park, FL 32789)

The present experiment was designed to determine if auditory imagery during reading includes talker-specific characteristics such as speaking rate. Following Kosslyn and Matt (1977), participants were familiarized with two talkers during a brief prerecorded conversation. One talker spoke at a fast speaking rate and one spoke at a slow speaking rate. During familiarization, participants were taught to identify each talker by name. At test, participants were asked to read two passages and told that either the slow or fast talker wrote each passage. In one condition, participants were asked to read each passage aloud, and in a second condition, they were asked to read each passage silently. Participants pressed a key when they had completed reading the passage, and reading times were collected. Reading times were significantly slower when participants thought they were reading a passage written by the slow talker than when reading a passage written by the fast talker. However, the effects of speaking rate were only present in the reading-aloud condition. Additional experiments were conducted to investigate the role of attention to talker's voice during familiarization. These results suggest that readers may engage in auditory imagery while reading that preserves perceptual details of an author's voice.

5aSC23. Duration as a cue in the distinction between glides and vowels. A. E. Coren (Dept. of Psych., Univ. of Texas, 1 University Station A8000, Austin, TX 78712-0187, aecoren@mail.utexas.edu) and W. J. Warren (Univ. of Texas, Austin, TX 78712-0198)

This experiment examines the role of length in the perception of glides and vowels. Tokens of the word "rodeo," spoken at two speech rates, were derived from an original by taking one cycle out of the waveform of the high vowel. Twenty listeners were presented with these derived tokens. It was found that when the length of the first vowel was approximately half that of the second, then 50 percent of the listeners judged the segment a vowel and 50 percent judged the segment a glide. This study agrees with earlier work done in this same manner. It is concluded that in running speech, the relative length of the first vowel is crucial in the vowel versus glide perceptual distinction.

5aSC24. Perceptual effects of preceding non-speech-rate information on temporal properties of speech categories. Travis Wade and Lori Holt (Dept. of Psych., Carnegie Mellon Univ., 5000 Forbes Ave., Pittsburgh, PA 15213, twade@andrew.cmu.edu)

The purpose of the study was to determine whether a general mechanism such as durational contrast can explain the apparent rate-dependent nature of speech perception. Four experiments were performed to test

effects of pure tone presentation rate on perception of following speech continua involving duration-varying formant transitions with which the tones shared critical temporal and spectral properties. Results showed small but consistent shifts in the stop-continuant boundary distinguishing /ba/ and /wa/ syllables based on the rate of precursor tones similar in duration and frequency to syllable-initial $F1$ and $F2$ transitions, across differences in amplitude of tones and despite variability in their duration. Additionally, the shift was shown to involve the entire graded structure of the [w] category and was not limited to an ambiguous boundary region, affecting goodness judgments on both sides of an estimated best exemplar range. These results are problematic for accounts of rate-dependent processing that explicitly reference speech categories or articulatory events and are consistent with a contrast account.

5aSC25. Medial surface dynamics of the vocal folds in an *in vivo* canine model. Michael Doellinger, Gerald S. Berke, Dinesh K. Chhetri, and David A. Berry (Div. of Head and Neck Surgery, UCLA School of Medicine, 31-24 Rehab., 1000 Veteran Ave., Los Angeles, CA 90095-1794)

Quantitative measurement of the medial surface dynamics of the vocal folds is important for understanding how sound is generated in the larynx. However, such data are hard to gather because of the inaccessibility of the vocal folds. Recent studies have applied hemi-larynx methodology to excised human larynges, to visualize these dynamics. The present study extends this methodology to obtain similar quantitative measurements using an *in vivo* canine hemi-larynx setup, with varying levels of stimulation to the recurrent laryngeal nerve. Use of an *in vivo* model allows us to examine effects of intrinsic muscle contraction on the medial surface of the vocal folds, to provide greater insight into mechanisms of vocal control. Data were collected using digital high-speed imaging with a sampling frequency of up to 4000 Hz, and a spatial resolution of up to 1024×1024 pixels. Three-dimensional motion will be extracted, computed, visualized, and contrasted as a function of the level of stimulation to the recurrent laryngeal nerve. Results will also be compared to patterns of vibration in excised larynges. Finally, commonly applied quantitative analyses will be performed to investigate the underlying modes of vibration. [Work supported by NIH/NIDCD.]

5aSC26. Perceptual relevance of source spectral slope measures. Jody Kreiman and Bruce R. Gerratt (Div. of Head/Neck Surgery, UCLA School of Medicine, 31-24 Rehab Ctr., Los Angeles, CA 90095-1794, jkreiman@ucla.edu)

Researchers broadly agree that the spectral slope of the voice source is an important concomitant of voice quality. Many measures of source spectral slope have been proposed, including the relative amplitudes of the lowest few harmonics, the ratio of energy in low- versus high-frequency bands, and the average deviation from an ideal slope. It is unclear which (if any) of these measures best reflects the differences in vocal quality that result from the underlying acoustic variability. To examine this issue, a large corpus of voice samples was inverse filtered, and spectra were calculated for resulting source pulses. Different measures of spectral slope were calculated for each voice, and correlations among measures were examined. Finally, several series of synthetic stimuli were created in which only the source spectral slope varied in steps. Listeners judged the similarity of stimuli within each series. Similarity responses were evaluated with multidimensional scaling, and the resulting perceptual spaces were interpreted in terms of the different measures of source spectral slope. Measures that are highly correlated with the perceptual spaces reflect perceptually important aspects of the source signal. [Research supported by NIDCD.]

5aSC27. Significance of analysis window size in maximum flow declination rate. Linda M. Carroll (Dept. of Otolaryngol., Mount Sinai School of Medicine, 5 E. 98th St., 1st Fl., Box 1653, Grabscheid Voice Ctr., New York, NY 10029, linda.carroll@mssm.edu)

Most acoustic stability measures and aerodynamic measures have recommended data analysis methods to ensure accurate comparison across research studies. Although standard aerodynamic analysis methods exist for subglottal pressure and mean transglottal flow, such standards do not exist for maximum flow declination rate (MFDR). As such, it becomes difficult to compare results across studies with the wide range of analysis windows. Because MFDR is a strong indicator of laryngeal function and increasingly is being reported in aerodynamic studies, it is necessary to determine whether a significant difference exists between analysis windows in data management. This study finds significant differences in MFDR with comparison of four different data extraction methods on the same data set. The contrasting data extractions compare differences between method A (MFDR from entire 1000-ms segment, excluding onset/offset), method B (MFDR from middle 100 ms with center at midportion of entire segment), method C (MFDR from middle 100 ms with center at midportion of maximum MFDR value from entire segment), and method D (MFDR from 20 consecutive cycles with center at midportion of maximum MFDR value from entire segment) for a sustained 7-syllable /pa/ repetition at two fundamental frequencies.

5aSC28. Fundamental frequency perturbation indicates perceived health and age in male and female speakers. David R. Feinberg (School of Psych., Univ. of St. Andrews, St. Andrews, Fife KY169JU, Scotland, UK)

There is strong support for the idea that healthy vocal chords are able to produce fundamental frequencies ($F0$) with minimal perturbation. Measures of $F0$ perturbation have been shown to discriminate pathological versus healthy populations. In addition to measuring vocal chord health, $F0$ perturbation is a correlate of real and perceived age. Here, the role of jitter (periodic variation in $F0$) and shimmer (periodic variation in amplitude of $F0$) in perceived health and age in a young adult (males aged 18–33, females aged 18–26), nondysphonic population was investigated. Voices were assessed for health and age by peer aged, opposite-sex raters. Jitter and shimmer were measured with Praat software (www.praat.org) using various algorithms (jitter: DDP, local, local absolute, PPQ5, and RAP; shimmer: DDA, local, local absolute, APQ3, APQ5, APQ11) to reduce measurement error, and to ascertain the robustness of the findings. Male and female voices were analyzed separately. In both sexes, ratings of health and age were significantly correlated. Measures of jitter and shimmer correlated negatively with perceived health, and positively with perceived age. Further analysis revealed that these effects were independent in male voices. Implications of this finding are that attributions of vocal health and age may reflect actual underlying condition.

5aSC29. Spectral anlaysis of glottal flow models. Matthew E. Lee (Ctr. for Signal and Image Processing, Georgia Inst. of Technol., mattlee@ece.gatech.edu) and Mark J. T. Smith (Purdue Univ.)

An analysis of the effects of time-domain characteristics of glottal waveforms in the spectral domain is presented. Studies have shown that the relative difference in amplitude of the first two harmonics of the inverse-filtered voice waveform ($H1^* - H2^*$) can be used to predict the value of the open quotient in certain situations [J. Sundberg, M. Andersson, and C. Hultqvist, J. Acoust. Soc. Am. **105**, 1965–1971 (1999)]. However, recent experiments have shown that glottal waveform asymmetry is an additional factor that will affect the relationship between the harmonics and open quotient. This work is aimed at establishing a method in which both the open quotient and a glottal asymmetry coefficient can be reliably determined based on the characteristics of the first two harmonics. By deriving analytic formulas for the spectra of two glottal flow models (LF, $R{+}{+}$), it is shown that a measure of the relative phase difference of the

5a FRI. AM

first two harmonics as well as the amplitude difference can be used in order to determine these perceptually important time-domain parameters of the glottal waveform. A theoretical analysis as well as experimental results are presented.

5aSC30. An improved correction formula for the estimation of voice source harmonic magnitudes. Markus Iseli (Dept. of Elec. Eng., UCLA, 405 Hilgard Ave., Los Angeles, CA 90095, iseli@ee.ucla.edu) and Abeer Alwan (UCLA, Los Angeles, CA 90095)

Information about voice and talker characteristics, or voice quality, can be gained from glottal flow estimates. Many voice quality parameters, such as the open quotient (OQ), depend on an accurate estimate of the glottal flow (voice source) spectrum. It is known that OQ, for example, is correlated with the magnitude difference of the first two harmonics (H_1-H_2) of the voice source spectrum [Holmberg *et al.*, "Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice," J. Speech Hear. Res. **38**, 1212−1223 (1995)]. To obtain an accurate estimate of the voice source spectral harmonics, the influence of vocal−tract resonances needs to be removed. In Hanson, Ph.D. dissertation, Harvard University (1995), a correction formula which removes the effect of F_1 on H_1 and H_2 was presented. In this work, an improved correction formula for the estimation of source harmonics′ magnitudes is presented. The new correction formula accounts for the bandwidths of vocal−tract resonances, and most importantly, is not limited to the analysis of nonhigh vowels. H_1-H_2 estimates, using the proposed technique with synthesized vowels /a/, /i/, and /u/ generated with the LF and the KLGLOTT88 models, are very accurate. [Work supported in part by the NSF.]

5aSC31. A comparison of harmonic production in trained and untrained singers. Jonathan H Hildebrand and Hilary J. Caso (Dept. of Linguist., Univ. of North Carolina, Chapel Hill, CB#3155, Chapel Hill, NC 27599, hildebr1@email.unc.edu)

This study was designed to investigate the differences in the tone qualities of trained and untrained singers of both genders by measuring the ratio of energy found in the fundamental frequency to that found in the harmonics of the two groups. Forty trained and untrained male and female subjects (ten in each group) sang the vowels [a e i o u] in the words "me cake father obey too." Each subject sang through the list of words in its entirety one time. They were also asked to produce each word at a constant pitch of their choosing. The intensity (dB) was then measured at the vowel midpoint in the intact vowel, and then again with the fundamental filtered out. These two measurements allowed for a comparison to be made by subtracting the results. This process produced results showing that the fundamental frequency was responsible for the majority of the trained singers', vocal productions, where harmonics were responsible for the majority of the untrained singers', vocal productions. The data also showed that these results were more prominent for the production of the high vowels in both genders, and for women overall.

5aSC32. Lexical frequency and voice assimilation in complex words in Dutch. Mirjam Ernestus, Mybeth Lahey, Femke Verhees, and Harald Baayen (Max-Planck Inst. for Psycholinguist. and Nijmegen Univ., Wundtlaan 1, 6525 XD, Nijmegen, The Netherlands, mirjam.ernestus@mpi.nl)

Words with higher token frequencies tend to have more reduced acoustic realizations than lower frequency words (e.g., Hay, 2000; Bybee, 2001; Jurafsky *et al.*, 2001). This study documents frequency effects for regressive voice assimilation (obstruents are voiced before voiced plosives) in Dutch morphologically complex words in the subcorpus of read-aloud novels in the corpus of spoken Dutch (Oostdijk *et al.*, 2002). As expected, the initial obstruent of the cluster tends to be absent more often as lexical frequency increases. More importantly, as frequency increases, the duration of vocal-fold vibration in the cluster decreases, and the duration of the bursts in the cluster increases, after partialing out cluster duration. This suggests that there is less voicing for higher-frequency words. In fact, phonetic transcriptions show regressive voice assimilation for only half of the words and progressive voice assimilation for one third. Interestingly, the progressive voice assimilation observed for higher-frequency complex words renders these complex words more similar to monomorphemic words: Dutch monomorphemic words typically contain voiceless obstruent clusters (Zonneveld, 1983). Such high-frequency complex words may therefore be less easily parsed into their constituent morphemes (cf. Hay, 2000), favoring whole word lexical access (Bertram *et al.*, 2000).

5aSC33. An acoustic and electroglottographic study of V-glottal stop-V in two indigenous American languages. Christina M. Esposito and Rebecca Scarborough (Dept. of Linguist., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095, esposito@humnet.ucla.edu)

Both Pima, a Uto-Aztecan language spoken in Arizona, and Santa Ana del Valle Zapotec (SADVZ), an Otomanguean language spoken in Oaxaca, Mexico, have sequences of two vowels separated by an intervening glottal stop. In both languages, this V?V sequence becomes reduced in certain occurrences, with the perceptual effect of the loss of /?/ in Pima and the loss of $V2$ in SADVZ. The purpose of this study is to provide an acoustic and electroglottographic (EGG) description of these sequences in both their full and reduced forms, prompted by varying speech rate. Two acoustic measures of phonation type ($H1-H2$, $H1-A3$) and two EGG measures (OQ and peak closing velocity) were made at vowel midpoints and adjacent to /?/. For Pima, an issue of interest is what properties of the /V?V/ sequences (when $V1=V2$) allow them to be distinguished from phonemic long vowels in the reduced forms where /?/ is lost. It is hypothesized that /?/ will be preserved as vowel glottalization. For SADVZ, an important question is why the vowels sound creaky despite a lack of spectral evidence for creak. It is hoped that more direct EGG measures will show the perceived phonation.

5aSC34. Acoustic and aerodynamic characteristics of ejectives in Amharic. Didier Demolin (Universidade de Sao Paulo/CNPq, 403 av. Prof. Gualberto, 01060-970 Sao Paulo, Brazil)

This paper invetsigates the main phonetic characteristics that distinguishes ejectives from pulmonic sounds in Amharic. In this language, there are five ejectives that can be phonemically singleton or geminate. Duration measurements have been made in intervocalic position for pulmonic stops and for each type of ejective, taking into account the overall duration and VOT. Results show that ejective stops have a higher amplitude burst than pulmonic stops. The duration of the noise is shorter for ejective fricatives compared to pulmonic fricatives. At the end of ejective fricatives, there is a 30-ms glottal lag that is not present in pulmonic fricatives. Geminate ejectives are realized by delaying the elevation of the larynx. This can be observed on the spectrographic data by an increase of the noise at the end of the geminate ejectives. Aerodynamic data have been collected in synchronization with the acoustic recordings. The main observations are that pharyngeal pressures values are much higher than what is usually assumed (up to 40 CmH2O for velars) and that the delayed command in the elevation of the larynx of geminate ejectives is shown by two phases in the rise of pharyngeal pressure.

5aSC35. Aerodynamic characteristics of French consonants. Didier Demolin (Universidade de Sao Paulo, 403 av. Prof. Gualberto, 01060-970 Sao Paulo, Brazil), Sergio Hassid (Hopital Ersame, Universite libre de Bruxelles), and Alain Soquet (Universite Libre de Bruxelles, 1050 Brussels, Belgium)

This paper reports some aerodynamic measurements made on French consonants with a group of ten speakers. Speakers were recorded while saying nonsense words in phrases such as papa, dis papa encore. The

nonsense words in the study combined each of the French consonants with three vowels /i, a, u/ to from two syllables words with the first syllable being the same as the second. In addition to the audio signal, recordings were made of the oral airflow, the pressure of the air in the pharynx above the vocal folds and the pressure of the air in the trachea just below the vocal folds. The pharyngeal pressure was recorded via a catheter (i.d. 5 mm) passed through the nose so that its open end could be seen in the pharynx below the uvula. The subglottal pressure was recorded via a tracheal puncture between the first and the second rings of the trachea or between the cricoid cartilage and the first tracheal ring. Results compare subglottal presssure, pharyngeal pressure, and airflow values. Comparisons are made between values obtained with male and female subjects and various types of consonants (voiced versus voiceless at the same place of articulation, stops, fricatives, and nasals).

5aSC36. Viscoelastic properties of the false vocal fold. Roger W. Chan (Otolaryngol. and Grad. Prog. in Biomed. Eng., Univ. of Texas Southwestern Medical Ctr., Dallas, TX 75390-9035, roger.chan@utsouthwestern.edu)

The biomechanical properties of vocal fold tissues have been the focus of many previous studies, as vocal fold viscoelasticity critically dictates the acoustics and biomechanics of phonation. However, not much is known about the viscoelastic response of the ventricular fold or false vocal fold. It has been shown both clinically and in computer simulations that the false vocal fold may contribute significantly to the aerodynamics and sound generation processes of human voice production, with or without flow-induced oscillation of the false fold. To better understand the potential role of the false fold in phonation, this paper reports some preliminary measurements on the linear and nonlinear viscoelastic behavior of false vocal fold tissues. Linear viscoelastic shear properties of human false fold tissue samples were measured by a high-frequency controlled-strain rheometer as a function of frequency, and passive uniaxial tensile stress-strain response of the tissue samples was measured by a muscle lever system as a function of strain and loading rate. Elastic moduli (Young's modulus and shear modulus) of the false fold tissues were calculated from the measured data. [Work supported by NIH.]

5aSC37. Determining the relationship between respiratory variables and fundamental frequency declination: A mathematical modeling of the data. Carole E. Gelfer (Dept. of Commun. Disord., William Paterson Univ., Wayne, NJ 07470) and Jay Jorgenson (City College of New York, New York, NY 10031)

Previous research [C. E. Gelfer, unpublished doctoral dissertation, City University of New York Graduate Center (1987)] demonstrated stability in fundamental frequency ($F0$) declination across uttererance of various lengths that appeared to be derived from a stable underlying subglottal pressure (Ps). Moreover, this stability was maintained even when airflow requirements were varied in response to the varying phonetic composition of utterances. However, although a preliminary mathematical analysis suggested that both the Ps and $F0$ could be characterized as second order linear systems, no formal testing of these curves was performed. In this study we compare these data employing the following statistical procedure. First, we hypothesize that the data from both sets can be modeled by a curve of the form $y = axbecx$, with unknown parameters a, b, and c which are computed. We then consider a multivariate confidence interval for the vector (a,b,c) of fitted parameters, as well as the difference of the vectors for two such models. The comparison is made through distributional estimates obtained via Monte Carlo simulation.

5aSC38. Flow characteristics in a scaled-up glottis model. Michael Barry, Timothy Wei (Dept. of Mech. and Aerosp. Eng., Rutgers Univ., Piscataway, NJ 08854), and Michael Krane (Rutgers Univ., Piscataway, NJ 08854)

Measurements of the velocity field of the flow through a scaled-up model of the vocal folds are presented. Scaling up the model ten times life size and using water as a working fluid enables temporally and spatially resolved measurements of the velocity using digital particle image velocimetry, making use of dynamic similarity. The Reynolds number (based on maximum glottal width and maximum jet speed) is 3000, and the Strouhal number (based on glottis length, maximum jet speed, and opening/closing time) ranges from 0.01 to 0.1. Prescribing the motion of the model vocal folds allows phase averaging of the flowfield. Measurements presented focus on the time intervals just after vocal fold opening and just before closure. [Work funded by NIDCD of NIH.]

5aSC39. Acoustic correlates of Japanese expressions associated with voice quality of male adults. Hiroshi Kido (Dept. of Commun. Eng., Tohoku Inst. of Technol., Sendai, Japan) and Hideki Kasuya (Utsunomiya Univ., Utsunomiya 321-8585, Japan)

Japanese expressions associated with the voice quality of male adults were extracted by a series of questionnaire surveys and statistical multivariate analysis. One hundred and thirty-seven Japanese expressions were collected through the first questionnaire and careful investigations of well-established Japanese dictionaries and articles. From the second questionnaire about familiarity with each of the expressions and synonymity that were addressed to 249 subjects, 25 expressions were extracted. The third questionnaire was about an evaluation of their own voice quality. By applying a statistical clustering method and a correlation analysis to the results of the questionnaires, eight bipolar expressions and one unipolar expression were obtained. They constituted high-pitched/low-pitched, masculine/feminine, hoarse/clear, calm/excited, powerful/weak, youthful/elderly, thick/thin, tense/lax, and nasal, respectively. Acoustic correlates of each of the eight bipolar expressions were extracted by means of perceptual evaluation experiments that were made with sentence utterances of 36 males and by a statistical decision tree method. They included an average of the fundamental frequency ($F0$) of the utterance, speaking rate, spectral tilt, formant frequency parameter, standard deviation of $F0$ values, and glottal noise, when SPL of each of the stimuli was maintained identical in the perceptual experiments.

5aSC40. Intrinsic fundamental frequency effects in hearing impaired speakers. Bryan Gick (Dept. of Linguist., Univ. of BC, 270-1866 Main Mall, Vancouver, BC V6T 1Z1, Canada), Barbara Bernhardt, and Penelope Bacsfalvi (Univ. of BC, Vancouver, BC)

The source of the well-known intrinsic fundamental frequency (IF0) effect of vowel height has been controversial for decades. Previous work has found the average IF0 effect cross-linguistically to be 15.3 Hz [Whalen and Levitt, J. Phonetics (1995)]. The present study investigates IF0 for four hearing-impaired speakers. Based on previous observations that profoundly hearing impaired speakers vary voice pitch less than normal hearing speakers [Osberger and McGarr, Speech Lang. (1982)], our participants were expected to show a reduced IF0 effect. However, results show an average IF0 effect of 22 Hz, with a markedly wide range across speakers, from -4 to 48 Hz, with three of the four participants showing an above average-sized effect. Further, results of measures taken following speech articulation intervention using visual feedback [Bernhardt *et al.*, Clin. Linguist Phonet. (2003)] show a decrease in IF0 for the speakers with an over-sized effect, and an increase in the speaker with an under-sized effect, despite that neither IF0 nor pitch in general were included in treatment. Results of this study support a lingual-articulatory origin for IF0, as well as suggesting that normal-hearing speakers may use auditory feedback to mediate what would otherwise be a larger effect. [Research supported by NSERC.]

5aSC41. The battle for $F0$: Glottalization versus stress in Statimcets. Sonya Bird and Marion Caldecott (Univ. of BC, 1866 Main Mall, Buchanan E270, Vancouver, BC V6T 1Z1, Canada)

The conflicts that arise in speech production give us insight into the interaction between faithfulness to mental representations and articulatory and acoustic limitations on speech. This paper presents an example of a perceptually based conflict [Kochetov, LabPhon8 (to appear)] in Statimcets, a Salish language spoken in the Interior of British Columbia. In a pilot study on glottalized resonants in Statimcets, Bird [ICSNL 38 (2003)] found that the primary cue to glottalization was creaky voicing, and that glottalization was reduced in stressed syllables. This paper expands on the previous study, incorporating new acoustic data on stress, and a more detailed acoustic analysis of glottalization. It is shown that one of the cues to stress, raised fundamental frequency ($F0$), conflicts with the lowered $F0$ associated with creaky voicing. In Statimcets, this conflict is resolved by preserving the perceptual salience of stress cues at the expense of cues to glottalization. This results in loss of glottalization in stressed syllables. Acoustic data and analyses will be presented. [Research supported by SSHRC.]

FRIDAY MORNING, 28 MAY 2004 CONFERENCE ROOM L, 8:55 TO 11:30 A.M.

Session 5aSP

Signal Processing in Acoustics: Selected Topics

Juliette Ioup, Chair

Physics Department, University of New Orleans, New Orleans, Louisiana 70148

Chair's Introduction—8:55

Contributed Papers

9:00

5aSP1. A multifilter approach to acoustic echo cancellation. John Usher, Wieslaw Woszczyk, and Jeremy Cooperstock (Ctr. for Interdisciplinary Res. in Music Media and Technol., McGill Univ., Montreal, QC, Canada, jusher@po-box.mcgill.ca)

Hands-free teleconferencing is increasingly frequent today. An important design consideration for any such communication tool that uses high-quality audio is the return echo caused by the acoustic coupling between the loudspeakers and microphones at each end of the conference. An echo-suppression filter (ESF) reduces the level of this return echo, increasing speech intelligibility. A new ESF has been designed based on a block frequency domain adaptive filter using the well-known least-mean-square (LMS) criteria. There are two important coefficients in LMS adaptive filters which affect how an ESF adapts to changing acoustic conditions at each end of the conference, such as double-talk conditions and moving electroacoustic transducers. Previous approaches to similar ESFs have used either a single or double pair of these coefficients, whereas the new model typically uses ten. The performance of single, double, and multifilter architectures was compared. Performance was evaluated using both empirical measurements and subjective listening tests. Speech and music were used as the stimuli for a two-way teleconferencing experiment. The new filter performed better than the single- and two-filter ESF designs, especially in conferencing conditions with frequent double talk, and the new ESF can be optimized to suit different acoustic situations.

9:15

5aSP2. An adaptive equalization scheme using acoustic energy density. Xi Chen, Scott D. Sommerfeldt, and Timothy W. Leishman (ESC N283, Brigham Young Univ., Provo, UT 84602, drcxchen@hotmail.com)

This paper presents an equalization system based on the filtered-x algorithm and acoustic energy density. Previous theoretical results have shown that equalization filters using acoustic energy density produce greater spatial uniformity in equalization than those using acoustic pressure. The paper discusses the implementation of the equalization system in a one-dimensional duct. An adaptive algorithm is utilized in the time domain to determine and apply the equalization filter. Results from a traditional filter design and the energy density filter design are compared.

9:30

5aSP3. Separation and identification of a transient signal using the characteristics of eigenvectors. James P. Larue (AFRL/IFEC, Rome, NY 13441), George B. Smith (Naval Res. Lab., Stennis Space Ctr., MS 39529), George E. Ioup, and Juliette W. Ioup (Univ. of New Orleans, LA 70148)

This is an extension of previous research [Larue *et al.*, J. Acoust. Soc. Am. **113**, 2212 (2003)] into the physical characteristics possessed by eigenvectors in a singular value decomposition (SVD) of a covariance matrix formed from a sinusoidal signal corrupted by multipath to which Gaussian noise is added. An animation of the SVD in progress will focus on the creation of the eigenvectors rather than the singular values. In this case, the SVD forms a three-part subspace decomposition corresponding to the three components of the signal. These subspaces are clearly determined from characteristics (Fourier transform and the Kaiser Varimax norm) associated with the eigenvectors and are not so easily determined from the singular values alone. [Research supported by the NRC-AFRL/IFEC and ONR.]

9:45

5aSP4. Continuous speech recognition using dynamic synapse neural network. Alireza A. Dibazar, Hassan H. Namarvar, and Theodore W. Berger (BME Dept., Univ. of Southern California, 3650 S. McClintock Ave., OHE 500, CA 90089-1451)

The modified architecture of the dynamic synapse neural network (DSNN) is used to model windowed short time speech signal. The quasi-linearization algorithm is applied to train the network. The parameters of the trained network, which are representatives of the signal, are fed into the GMM/HMM based classifier. The performance of the modified architecture with GMM/HMM based classifier is demonstrated by recognition of continuous speech from unprocessed, noisy raw waveforms spoken by multiple speakers. Our results indicate that the features obtained from DSNN are robust in the presence of additive white Gaussian noise with respect to state-of-the-art Mel frequency features. [Work supported in part by DARPA, NASA and ONR.]

10:00–10:15 Break

10:15

5aSP5. A sound-texture detection algorithm. Michael J. Norris and Susan L. Denham (Ctr. for Theoretical and Computational Neurosci., Univ. of Plymouth, PL1 8AA, UK, mjnorris@plymouth.ac.uk)

A method for comparing sound textures based on the estimation of high-order information redundancy is presented. To calculate redundancy, a histogram is constructed that records the probability of occurrence of each possible short sequence of (typically about eight) samples. A self-organizing map [T. Kohonen, Proc. IEEE **78**, 1464–1480 (1990)] is used to summarize the histogram. The resulting probability distribution, or a trained self-organizing map that represents it, captures the texture of the sound, and leads to a natural definition of sound texture in terms of statistical stationarity. Sound textures such as recordings of running water, cafeteria noise, and traffic noise can be intuitively distinguished and classified by this method, despite small audible variations in amplitude, frequency, filtering, noise, and sound mixtures. The algorithm requires no preprocessing, Fourier analysis, heuristics, or finely tuned parameters, and may be implemented using a simple connectionist architecture. Textures can be compared through time by decaying probability distributions as new information arrives. A mechanism based on redundancy through time can also detect arbitrary changes in a sound, providing a novel mechanism to simulate auditory change detection. This versatile algorithm may find diverse applications in signal processing and neural computation. [Work supported by EU Open FET IST-2001-38099.]

10:30

5aSP6. Psychoacoustics based gain compensation for low listening levels. Thorvaldur Einarsson and Carol Espy-Wilson (Dept. of Elec. and Computer Eng., Univ. of Maryland, College Park, MD 20742)

Although the human hearing system is very complex and only understood to a limited degree, many measurements and models exist that explain parts of the hearing system. This paper uses one of these models, the contours of equal loudness, along with DSP techniques, to make music played at low listening levels sound more like it does at the intended listening level. The perceived frequency balance of music varies with listening level. This is especially noticeable at low listening levels, where frequencies below 500 Hz seem attenuated. Moreover, hearing perception exhibits nonlinear dynamic range compression, most evident at low frequencies. Traditional methods add fixed low-frequency gain to compensate for perceptual low-frequency attenuation at low levels. These methods do not consider the dynamic range compression and its nonlinearity and are often characterized by a boomy, unnatural sound. A system is designed where filter banks and power measurements estimate the time-varying power of low-frequency parts of the audio signal. The time-varying power of each narrow frequency band is compared to the contours of equal loudness, and changes are made to get the same frequency balance as at the intended listening level. The paper covers the design, implementation and performance of this system.

10:45

5aSP7. Two-dimensional Gabor analysis of space–time transient Lamb waves using Laser ultrasonic investigation. Loic Martinez (IUPGE ECIME UCP, 5 mail Gay Lussac, 95031 Cergy Pontoise, France, loic.martinez@iupge.u-cergy.fr), Nikolaas Van Riet, and Christ Glorieux (Katholieke Universiteit, Leuven, B-3001 Leuven, Belgium)

Laser generation/detection methods allow the investigation of ultrasonic transient phenomena in both space and time dimensions. Used for the experimental investigation of surface wave propagation along a one-dimensional medium, laser ultrasonics leads to two-dimensional (2D) space–time signal collections. In order to extract the wave propagation information, the classical high resolution signal processing methods or 2D Fourier transforms are not suitable to identify the transient and local aspect of wave propagation and mode conversion. In order to quantify these transient aspects in the space–time-wave number–frequency domains, the 2D Gabor transform is introduced. Its potential for the identification of the local and transient complex wave numbers is illustrated on the propagation of Lamb waves on a plane limited plate. The mode conversion sequences are clearly identified in the space–time–wave number-frequency planes. The experimental results are in good agreement whith the numerical simulations.

11:00

5aSP8. Phase velocimetry based on the spatio-temporal gradient analysis for detecting subsurface defects. Kenbu Teramoto (Dept. of Mech. Eng., SAGA Univ., Saga 8408502, Japan, tera@me.saga-u.ac.jp)

Ultrasonic Lamb-wave techniques are potential candidates of nondestructive evaluation (NDE) methodology: they allow the detection of surface defects or internal delamination. Because they propagate over a long distance, weakly attenuated in the case of a free-elastic plate, they are used for the evaluation of the assembly of large free plates in aeronautics, for example. The phase velocity of A0-mode Lamb-wave increases with thickness of the plate. Delaminations or internal defects, therefore, make the phase velocity slower. In this paper, an instantaneous phase velocimetry based on the spatio-temporal gradient analysis is proposed. The proposed method has an ability to measure the phase velocity of A0-mode Lamb wave field through the linearity among the four-dimensional vectors which is defined by following components: (1) a vertical displacement; (2) its vertical velocity; and (3) and (4) a pair of out-of-plane shear strains of the plate. In this study, the computational process of the local velocimetry is discussed and its physical meanings are investigated through FDTD simulations. Several results obtained through the proof-of-concept model are in relatively good agreement with the simulations.

11:15

5aSP9. Implementation of algorithms for extracting tonal components in underwater noise. M. H. Supriya and P. R. Saseendran Pillai (Dept. of Electron., Cochin Univ. of Sci. and Technol., Cochin-22, Kerala, India)

Underwater target classification and tracking problems utilize the noise signals emanating from the targets as well as the target dynamics features. Target classification is carried out by extracting certain classification clues about the target such as its emission frequencies, and other target specific features from the self noise or active transmissions. The self noise generated by the targets being nonstationary in nature are to be analyzed with short term averaged data segments for extracting the tonal components using various spectral estimation techniques. Though classical spectral analysis techniques are computationally efficient, it suffers from several inherent limitations such as frequency resolution, performance degradation due to implicit windowing of the data, etc. In an attempt to improve the spectral resolution, several modern spectral estimation techniques, which utilize the procedures of indirect Fourier analysis by fitting the measured short data segments to an assumed model, have been evolved. This paper presents the development of procedures for the estimation of power spectral densities using modern techniques based on parameter estimation such as auto regressive, moving average, auto regressive moving average, maximum likelihood, etc. The performance of the estimator has been validated by computing the emission frequencies of a 50-foot vessel.

Session 5aUW

Underwater Acoustics and Engineering Acoustics: Autonomous Underwater Vehicle Acoustics: Part I: Concepts and Systems

Henrik Schmidt, Cochair
Department of Ocean Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Thomas R. Howarth, Cochair
NAVSEA Newport, 1176 Howell Street, Newport, Rhode Island 02841

Chair's Introduction—8:35

Invited Papers

8:40

5aUW1. Dolphin sonar detection and discrimination capabilities. Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734)

Dolphins have a very sophisticated short range sonar that surpasses all technological sonar in its capabilities to perform complex target discrimination and recognition tasks. The system that the U.S. Navy has for detecting mines buried under ocean sediment is one that uses Atlantic bottlenose dolphins. However, close examination of the dolphin sonar system will reveal that the dolphin acoustic hardware is fairly ordinary and not very special. The transmitted signals have peak-to-peak amplitudes as high as 225–228 dB *re* 1 μPa which translates to an rms value of approximately 210–213 dB. The transmit beamwidth is fairly broad at about 10o in both the horizontal and vertical planes and the receiving beamwidth is slightly broader by several degrees. The auditory filters are not very narrow with Q values of about 8.4. Despite these fairly ordinary features of the acoustic system, these animals still demonstrate very unusual and astonishing capabilities. Some of the capabilities of the dolphin sonar system will be presented and the reasons for their keen sonar capabilities will be discussed. Important features of their sonar include the broadband clicklike signals used, adaptive sonar search capabilities and large dynamic range of its auditory system.

9:00

5aUW2. AUVs as integrated, adaptive acoustic sensors for ocean exploration. Henrik Schmidt, Joseph R. Edwards, Te-Chih Liu, and Monica Montanari (Dept. of Ocean Eng., MIT, Cambridge, MA 02139, henrik@mit.edu)

Autonomous underwater vehicles (AUV) are rapidly being transitioned into operational systems for national defense, offshore exploration, and ocean science. AUVs provide excellent sensor platform control, allowing for, e.g., accurate acoustic mapping of seabeds not easily reached by conventional platforms, such as the deep ocean. However, the full potential of the robotic platforms is far from exhausted by such applications. Thus, for example, most seabed-mapping applications use imaging sonar technology, the data volume of which cannot be transmitted back to the operators in real time due to the severe bandwidth limitation of the acoustic communication. The sampling patterns are therefore in general being preprogramed and the data are being stored for postmission analysis. This procedure is therefore associated with indiscriminate distribution of the sampling throughout the area of interest, irrespective of whether features of interest are present or not. However, today's computing technology allows for a significant amount of signal processing and analysis to be performed on the platforms, where the results may then be used for real-time adaptive sampling to optimally concentrate the sampling in area of interest, and compress the results to a few parameters which may be transmitted back to the operators. Such adaptive sensing concepts combining environmental acoustics, signal processing, and robotics are currently being developed for concurrent detection, localization, and classification of buried objects, with application to littoral mine countermeasures, deep ocean seabed characterization, and marine archeology. [Work supported by ONR and NATO Undersea Research Center.]

9:20

5aUW3. New virtual sonar and wireless sensor system concepts. B. H. Houston, J. A. Bucaro, and A. J. Romano (Naval Res. Lab., 4555 Overlook Ave., SW Washington, DC 23075)

Recently, exciting new sensor array concepts have been proposed which, if realized, could revolutionize how we approach surface mounted acoustic sensor systems for underwater vehicles. Two such schemes are so-called ''virtual sonar'' which is formulated around Helmholtz integral processing and ''wireless'' systems which transfer sensor information through radiated RF signals. The ''virtual sonar'' concept provides an interesting framework through which to combat the dilatory effects of the structure on surface mounted sensor systems including structure-borne vibration and variations in structure-backing impedance. The ''wireless'' concept would eliminate the necessity of a complex wiring or fiber-optic external network while minimizing vehicle penetrations. Such systems,

however, would require a number of advances in sensor and RF waveguide technologies. In this presentation, we will discuss those sensor and sensor-related developments which are desired or required in order to make practical such new sensor system concepts, and we will present several underwater applications from the perspective of exploiting these new sonar concepts. [Work supported by ONR.]

9:40

5aUW4. Bluefin autonomous underwater vehicles: Programs, systems, and acoustic issues. Joseph E. Bondaryk (Bluefin Robotics Corp., 301 Massachusetts Ave., Cambridge, MA 02139, bondaryk@bluefinrobotics.com)

Bluefin Robotics Corporation has been manufacturing autonomous underwater vehicles (AUVs) since spinning out of the MIT Sea Grant Laboratory in 1997. Bluefin currently makes three different diameter models of AUVs; the 9, 12, and 21, all based on the same free-flooded architecture and vectored-thrust propulsion design. Auxiliary acoustic systems include acoustic abort, ranging beacons, and acoustic modems. Vehicle navigation is aided by a downward-looking acoustic Doppler velocity logger (DVL). Sonar payloads can include: bottom profiler, side-scan sonar, SAS, forward-looking imagers (DIDSON), as well as horizontal and vertical discrete hydrophone arrays. Acoustic issues that arise include: (1) transmission of sound through the ABS plastic vehicle shell; (2) the impact of vehicle self-noise on data; (3) interoperability of sonars with other acoustic emitters present on and off the vehicle; and (4) the impact of navigation on some acoustic operations like SAS. This talk will illustrate these issues with real data collected on various Bluefin vehicles.

10:00–10:20 Break

10:20

5aUW5. AUV-based synthetic aperture sonar: Initial experiences and insights. Daniel A. Cook, Jose E. Fernandez, John S. Stroud, Kerry W. Commander, and Anthony D. Matthews (Naval Surface Warfare Ctr., Panama City, Code R21, Panama City, FL 32407-7001)

The ability to do synthetic aperture sonar (SAS) imaging from autonomous underwater vehicles (AUVs) has only recently been achieved. The combination of the two technologies is a milestone in the field of underwater sensing as the combination of high-resolution SAS imaging with AUVs will provide military, research, and commercial users with systems of unprecedented performance and capabilities. The U.S. Navy took delivery of the first AUV-based SAS in early 2003. A description of the system will be presented along with the methodologies employed. The emphasis will be on image quality and repeatability, as well as the differences associated with operating an AUV-based SAS as opposed to a towed SAS. Additional topics will include general comments and recommendations for better AUV/SAS integration such as mission planning and vehicle control strategies intended to maximize the chances of high-quality imagery, SAS motion measurement requirements coupled with on-board vehicle navigation, and the potential of using the SAS data to augment the vehicle motion sensors. Lastly, a brief overview of forthcoming Navy SAS systems will be included.

10:40

5aUW6. Passive acoustic localization with an AUV-mounted hydrophone array. Gerald L. D'Spain, Eric Terrill, C. David Chadwell, Jerome A. Smith, and Richard Zimmerman (Marine Phys. Lab., Scripps Inst. of Oceanogr., La Jolla, CA 93940-0704)

A mid-size Odyssey IIb autonomous underwater vehicle (AUV) was retrofitted with the advanced vectored-thrust system presently installed on AUVs manufactured by Bluefin Robotics, Inc. Subsequent modifications to this thrust system decreased the radiated acoustic and vibration noise levels recorded by an eight-element hydrophone array mounted on the AUV's inner shroud by 20 to 50 dB across the 20 Hz to 10 kHz band. This reduction in self-noise levels to near, or at, background ocean noise levels permits the use of the vehicle-mounted hydrophone array in passive ocean acoustic studies. One example is the application of passive synthetic aperture processing techniques to provide greater spatial resolution estimates of the direction of low frequency sources. Doppler spreading caused by medium motion is a limiting factor in array gain. At mid frequencies (1–10 kHz), the complexity of the received acoustic field created by scattering off the AUV body is partly captured in the array processing by the use of replica vectors measured in a calibration tank. These empirical replica vectors decrease the azimuthally dependent degradation in beamforming performance over that of plane waves. [Work supported by ONR, Code 321(US).]

Contributed Papers

11:00

5aUW7. Exploiting autonomous underwater vehicle mobility for enhanced sonar performance. Joseph R. Edwards and Henrik Schmidt (MIT Dept. of Ocean Eng., 77 Massachusetts Ave., Rm. 5-204, Cambridge, MA 02139)

Autonomous underwater vehicles (AUVs) provide a mobile sonar platform for local environmental characterization and mine hunting missions in areas that are prohibitively dangerous or otherwise inaccessible to manned vessels. Such missions are typically implemented by providing the vehicle with a series of way-points that uniformly sample the region of interest. While these preplanned sampling paths can be effective for many purposes, a higher degree of efficiency and efficacy can be achieved when the AUV is able to react to its perceived environment. Such response to the sensory stimulus represents a preliminary step toward the AUV achieving tasks in the way that a dolphin or human might. Upon sensing an object or feature of interest, the AUV can further interrogate the object by obtaining multiple views from preferred vantage points. Probabilities of detection and correct classification can be greatly impacted by a wise choice of sonar-adaptive AUV behavior. In this paper, the ability of the AUV to improve sonar performance through sonar-adaptive mission planning is investigated in the context of the mine hunting problem. Trade-offs

are discussed for continuous versus scripted mission adaptations, single-minded versus multiple-objective mission control, and task-wise versus mission-wise behavior. [Work supported by ONR and SACLANTCEN.]

11:15

5aUW8. Surface ship wake remote sensing using an AUV-mounted multibeam sonar. Lee Culver and David Bradley (Appl. Res. Lab., The Penn State Univ., P.O. Box 30, State College, PA 16804)

A 250-kHz multibeam sonar integrated into a high-speed autonomous underwater vehicle (AUV) has been used to image the wakes of two large surface ships from the underside. Using vehicle autopilot data (yaw, pitch, and roll) and acoustic tracking data, the multibeam sonar data were processed to produce 3D maps of the ship wake intensity. Also, the backscattered signal has been used to estimate the number of bubbles per unit volume (the bubble density), assuming a particular size distribution of the bubbles, using a technique that other researchers have employed to estimate the bubble density from single beam sonar data. These measurements have been discussed in two earlier ASA talks [J. Acoust. Soc. Am. **110**, Pt. 2 (2001); **111**, Pt. 2 (2002)]. The present talk focuses on how vehicle motion is used to process acoustic data from individual beams so as to produce a 3D map of the ship wake. Future plans involving a commercial multibeam sonar integrated into a high speed vehicle are discussed. [Work sponsored by NAVSEA PMS 415 and ONR Code 333.]

11:30

5aUW9. Concurrent detection and classification of targets with multistage signal-processing algorithms. Monica Montanari, Joseph R. Edwards, and Henrik Schmidt (Dept of Ocean Eng., MIT, 77 Massachusetts Ave., Rm. 5-204, Cambridge, MA 02139, momo@mit.edu)

Concurrent detection and classification (CDAC) of targets stands as the goal in littoral mine-hunting missions. CDAC systems commonly apply model-based algorithms that include *a priori* known features of the target inside the detection algorithm. If the models are accurate, then this approach significantly reduces the false-alarm rate inherent in detection-only methods. When the possible targets are unknown, as may be the case in tactical situations, then these model-based methods not only fail to reduce the false-alarm rate, but may also reduce the probability of detection. Simultaneous optimization of detection and classification presents a challenge due to competing criteria; detection seeks to integrate signals to improve signal-to-noise ratio, while classification seeks to preserve small features of distinction within the signals. In this work, a method for robust CDAC is demonstrated that exploits the capabilities of autonomous underwater vehicles (AUVs) and multistage signal-processing algorithms to systematically investigate targets of interest in a single mission. The deformable geometry of the AUV-borne sonar network is exploited to provide favorable views of targets to achieve multiple objectives in series, and on-board computational facilities allow the implementation of multiple signal-processing regimes [Work supported by ONR and NATO Undersea Research Centre.]

11:45

5aUW10. Intervessel navigation using range and range rate. Brian S. Bourgeois and Patrick M. McDowell (Naval Res. Lab., Stennis Space Ctr., MS 39529, bsb@nrlssc.navy.mil)

A fundamental requirement for groups of unmanned underwater vehicles to work cooperatively together is the ability for each vessel to know the relative position of its neighbors. While external communication and positioning infrastructures can be used for this purpose, a more flexible approach is to give each vessel the ability to independently discern the location of its neighbors using its own communication and sensor systems. Recent developments in acoustic modems have included the ability to measure range between two vessels as well as the Doppler imparted on the signal by the relative motion between vessels. The measured Doppler shift can be used to compute the instantaneous range rate between two vessels; researchers at NRL are presently working to validate these measurements using GPS. Approaches to vessel relative positioning and navigation using successive range and range-rate measurements, along with directed vessel maneuvers, will be presented.

FRIDAY AFTERNOON, 28 MAY 2004 — NEW YORK BALLROOM A, 1:00 TO 2:30 P.M.

Session 5pAOa

Acoustical Oceanography: Ocean Basin Acoustics

John L. Spiesberger, Chair
Department of Earth and Environmental Science, University of Pennsylvania, 158 Hayden Hall, 240 South 33rd Street, Philadelphia, Pennsylvania 19104-6316

Contributed Papers

1:00

5pAOa1. Acoustic identification of a single transmission at 3115 km from a bottom-mounted source at Kauai. John L. Spiesberger (Dept. of Earth and Environ. Sci., 240 S. 33rd St., Univ. of Pennsylvania, Philadelphia, PA 19104, johnsr@sas.upenn.edu)

Sounds received in the Gulf of Alaska at 3115 km from the ATOC/NPAL source at Kauai (75-Hz, 0.027-s resolution, bottom-mounted) are compared with acoustic and oceanographic models. Unlike data collected at stationary SOSUS arrays, these data come from a towed horizontal array at 372-m depth of military origin. A plausible identification of the acoustic reception is made despite the fact that only one transmission is collected and sound interacts with the bottom near the source. The similarity between the modeled and measured impulse response here may be useful for understanding the signals between this same source and the NPAL array near southern California. The plausible identification of sound from the horizontal array here appears to point toward the feasibility of using other military platforms of opportunity besides SOSUS to study acoustic propagation and possibly map climatic changes in temperature by means of tomography.

1:15

5pAOa2. U.S. Navy sources and receivers for studying acoustic propagation and climate change in the ocean. John L. Spiesberger (Dept. of Earth and Environ. Sci., 240 S. 33rd St., Univ. of Pennsylvania, Philadelphia, PA 19104, johnsr@sas.upenn.edu)

Sounds from a U.S. Navy SSQ-110A source are received at high signal-to-noise ratios at ocean-basin scales at two Sound Surveillance Systems in the Pacific. The sounds have sufficient pulse resolution to study climatic variations of temperature. The acoustic data can be understood

using ray and parabolic approximations to the wave equation. Modeled internal waves decrease pulse resolution from 0.01 to 0.1 s, consistent with observations.

1:30

5pAOa3. Science enabled by ocean observatory acoustics: The NSF ORION program. Bruce M. Howe (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698, howe@apl.washington.edu), James H. Miller (Univ. of Rhode Island, Narragansett, RI 02882), and IASOO Committee (Integrated Acoust. Systems for Ocean Observatories Committee, Acoust. Society of America)

The National Science Foundation (NSF) has started the Ocean Research Interactive Observatories Network (ORION) program for research-driven sustained observations. The core infrastructure will consist of: (1) a coarse global array of buoys, (2) a regional cabled observatory in the northeast Pacific (with Canada already funded for the northern portion), and (3) coastal observatories. Seafloor junction boxes providing power and communications are a common enabling feature. The ORION Workshop (Puerto Rico, 4–8 January 2004) developed science themes that can be addressed utilizing this infrastructure. Acoustics enable much of the science. The use of acoustics to sense the synoptic 3-D/volumetric ocean environment was found to be ubiquitous through most ORION working groups. One reason for this is the relative transparency of the ocean to sound and the opaqueness to electromagnetic radiation. Participants at the workshop formed an Acoustics Working Group. Based on its report, we review the science and technical drivers for acoustics and educational opportunities. Themes include inherent volumetric, near instantaneous sampling, robust transducers, imaging at many scales, navigation, communications, and using sound in the sea as a major education and outreach mechanism. Recommendations include the formation of a standing ORION committee on acoustics and a workshop. See http://www.orionprogram.org and http://www.oce.uri.edu/ao/AOWEBPAGE.

1:45

5pAOa4. Horizontal coherence in the NPAL experiment. Michael Vera, Matthew Dzieciuch (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0225), and The NPAL Group[a)] (APL-UW, SIO-UCSD, WHOI)

Acoustic transmissions from a broadband source near Kauai with a center frequency of 75 Hz were recorded on a two-dimensional receiver array at a range of 3900 km as part of the North Pacific Acoustic Laboratory (NPAL) experiment. The receiver array consisted of five vertical line arrays (VLAs), with separations transverse to the propagation path ranging from a few hundred meters to a few kilometers. The coherences derived from the data have been compared to two different numerical predictions based on sound-speed perturbations due to internal waves. An approximation to the acoustic path integral yields a prediction for the length scale of horizontal coherence. Parabolic-equation simulations of propagation from the source to each VLA through multiple realizations of a random internal-wave field provide another estimate of the coherence at each horizontal separation. [Work supported by ONR.] [a)]J. A. Colosi, B. D. Cornuelle, B. D. Dushaw, M. A. Dzieciuch, B. M. Howe, J. A. Mercer, R. C. Spindel, and P. F. Worcester.

2:00

5pAOa5. Precise measurement of travel time in 1999 OAT experiment. Wang Yong (Grad. of Sci. & Technol., Chiba Univ., 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan, wang@graduate.chiba-u.jp) and Hachiya Hiroyuki (Chiba Univ., Chiba 263-8522, Japan)

The 1999 OAT experiment is performed in the Central Equatorial Pacific to monitor the ocean phenomenon associated with El Niño and the Southern Oscillation (ENSO). In order to analyze the travel-time perturbations due to ocean current, amplitude information of the received signal is used in the conventional method. However, it is difficult to measure the travel-time difference precisely, mainly because of the attenuation of long-range transmission. In this report, we propose a new technique which is the complex vector method by using the amplitude and phase information. From the ray identification results, we estimated the travel-time difference between corresponding rays of the reciprocal transmissions. Since the phase of a ray signal is very stable in 130-s signal duration time, the precision of the measurement of the travel-time difference using phase difference is higher than the conventional estimation. The travel-time differences estimated by using this method have reasonable magnitudes.

2:15

5pAOa6. Sensitivity kernels of finite-frequency travel times in ocean acoustic tomography. Emmanuel K. Skarsoulis (Inst. of Appl. and Comput. Math. FORTH, 711 10 Heraklion, Crete, Greece, eskars@iacm.forth.gr) and Bruce D. Cornuelle (Scripps Inst. of Oceanogr., La Jolla, CA 92093)

Wave theoretic modeling is applied to obtain travel-time sensitivity kernels representing the amount by which travel times are affected by localized sound-speed variations anywhere in the medium. In the ray approximation travel times are sensitive to medium changes only along the corresponding eigenrays. In the wave-theoretic approach the perturbations of peak arrival times are expressed in terms of pressure perturbations, which are further related with the underlying sound-speed perturbations using the first Born approximation. In this way, an integral representation of travel-time perturbations is obtained in terms of sound-speed perturbations; the associated kernel represents the spatial sensitivity of travel times to sound-speed perturbations. The application of the travel-time sensitivity kernel to an ocean acoustic waveguide gives a picture close to the ray-theoretic one in the high-frequency case but significantly differs at lower frequencies. Low-frequency travel times are sensitive to sound-speed changes in areas surrounding the eigenrays, but not on the eigenrays themselves, where the sensitivity is zero. Further, there are areas of positive sensitivity, where, e.g., a sound-speed increase results in a counter-intuitive increase of arrival times. These findings are confirmed by independent forward calculations.

Session 5pAOb

Acoustical Oceanography: Passive and Active Observations

Nicholas C. Makris, Chair
Department of Ocean Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Contributed Papers

2:45

5pAOb1. Prediction of ocean ambient sound generated from geophysical signals. Barry Ma, Jeffrey Nystuen, and Ren-Chieh Lien (APL, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698)

About 90 buoy months of ocean ambient sound data are collected using Acoustic Rain Gauges in different open-ocean locations. Distinct ambient sound spectra are identified through a series of discrimination process. Some distinct features are described as follows: (1) Excluding effects of rain, observed ambient sound spectra, generated exclusively by wind, exhibit a constant frequency spectral slope in 1–35 kHz. (2) Drizzle produces a prominent spectral peak at 15 kHz, and the magnitude of the spectral peak is very sensitive to the wind speed. These are consistent with earlier findings [Vagle *et al.* (1990) and Nystuen (1993)]. (3) Preliminary analysis shows that the spectral slope between 1–10 kHz decreases linearly with the rainfall rate. Therefore, the rainfall rate may be reliably predicted by both the spectral slope and the spectral magnitude in 1–10 kHz frequency range. The comparison of these two independent predictions provides the quality check of the estimate of the rainfall rate. An analytical spectral model of the ocean ambient sound from 1–50 kHz is constructed using two input parameters, rainfall rate and wind speed, based on the existing observations. This analytical model is aimed at predicting the ambient sound spectra at varying rainfall rates and wind speeds.

3:00

5pAOb2. A novel technique for measuring the rainfall from the rain noise. Mani Thundiyil and P. R. Saseendran Pillai (Dept. of Electron., CUSAT, Cochin 682022, India)

It has been observed that if water droplets are allowed to fall with terminal velocity, on to the water trapped in the chamber of a specially designed sensor assembly, the resulting acoustic signal generated from the drop impact is wideband and extends up to 100 kHz. The acoustic energy of the signal in the low-frequency range is seen to be vis-a-vis correlated to the kinetic energy of the drops. Such an assembly is exposed to the rain to pick up the rain-generated noise. The energy of the rain-generated noise is seen to be proportional to the kinetic energy of the rain. From the rain kinetic energy–intensity relationship, the rainfall rate is computed. This paper presents an experimental procedure for estimating the rain intensity from the rain generated underwater noise captured by the sensor assembly. Experiment has been carried out for various rainfall rates and the computed rain intensities were compared with the ones measured with a tipping bucket rain gauge. The results and inferences suggest that this approach is a simple, cost effective and computationally efficient technique to measure the kinetic energy of rain as well as compute the rainfall rate.

3:15

5pAOb3. Experimental demonstration of accurate hurricane classification from local wind speed estimates obtained with underwater sound. Joshua D. Wilson and Nicholas C. Makris (77 Massachusetts Ave. 5-212, Cambridge, MA 02139, makris@mit.edu)

In 1999 Hurricane Gert passed over an autonomous hydrophone in the North Atlantic yielding a clear recording of the underwater noise generated by the hurricane. By correlating the noise with meteorological data from reconnaissance aircraft and satellites, it is shown that underwater noise intensity in the 10–50-Hz band is proportional to local wind speed to the 3.3 power. Previous experiments have shown similar relationships at lower wind speeds in many different ocean environments [Crouch, J. Acoust. Soc. Am. **51**, 1066–1072 (1972); Marrett and Chapman, IEEE J. Ocean. Eng. **15**, 311–315 (1990); Chapman and Cornish, J. Acoust. Soc. Am. **93**, 782–789 (1993)]. Given this empirical relationship, it is shown that acoustic intensity measurements can be used to accurately estimate wind speed to within a 5% error margin, and so can be used to accurately classify the destructive power of a hurricane. Hurricane Gert, for example, is found to be a class 2 hurricane with maximum wind speeds of at least 89 kts by this underwater acoustic approach, which is in accord with aircraft measurements. Potential advantages of this underwater acoustic approach are discussed as are their potential impact on hurricane forecasting and disaster planning.

3:30

5pAOb4. Explaining extended linear features observed in remote sonar images of the New Jersey continental shelf break during Acoustic Clutter Experiments in 2001 and 2003. Sunwoong Lee, Purnima Ratilal, and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

Prominent acoustic clutter forming a lineated feature spanning more than 10 km has been repeatedly observed in long-range sonar imagery acquired during the Acoustic Clutter Experiments in 2001 and 2003. The lineated feature appears on the continental shelf near the shelf break for downslope sonar positions. It is hypothesized that this lineated feature is caused by scattering from fish shoals, seafloor slopes, or subbottom geology such as the R-reflector. A range-dependent reverberation and scattering model using the parabolic equation [Ratilal *et al.*, J. Acoust. Soc. Am. **114**, 2302 (2003)] is employed to test these various hypotheses by model–data comparison using high resolution bathymetry, subbottom profiles, and fish density data acquired in support of the long-range acoustic measurements.

3:45

5pAOb5. Inferring fish school distributions from long range acoustic images: Main acoustic clutter experiment 2003. Deanelle T. Symonds, Purnima Ratilal, Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139), and Redwood W. Nero (Naval Res. Lab, Stennis Space Ctr., MS 39529)

Long range scattering from fish schools and bottom reverberation in the New Jersey Continental Shelf environment are modeled using a unified, range-dependent, bistatic scattering, and reverberation model based on the parabolic equation [Ratilal and Makris, J. Acoust. Soc. Am. **114**, 2302 (2003)]. The fish swim bladder is approximated as an air-filled bubble, while the bottom reverberation from volume inhomogeneities is modeled using the Rayleigh–Born approximation. The broadband scattered field, in the frequency range from 390 to 440 Hz, is beamformed and spatially charted using two-way travel time. The model output is compared with scattered field levels from fish schools and background reverberation measured during the the Main Acoustic Clutter Experiment 2003 using a long range, bistatic sonar system. The fish school characteristics, such as size, distribution and density, are inputs to the model. These are obtained from measurements made by the fish finding sonar during the experiment. This calibrated model is then used to infer fish school distributions and densities in areas where fish finding sonar measurements are not available.

4:00

5pAOb6. Continuous wide area monitoring of fish shoaling behavior with acoustic waveguide sensing and bioclutter implications. Nicholas C. Makris, Purnima Ratilal, Deanelle T. Symonds (MIT, 77 Massachusetts Ave., Cambridge, MA 02139, makris@mit.edu), and Redwood W. Nero (Naval Res. Lab, Stennis Space Ctr., MS 39529)

Field measurements are used to show that the detailed behavior of fish shoals can be continuously monitored at roughly 1-min intervals over wide areas spanning hundreds of square kilometers by long range acoustic waveguide sensing. The technique was used on the New Jersey Continental Shelf to produce unprecedented video images of shoal formation, fragmentation, and migration. Simultaneous line-transect measurements show the imaged shoals to contain pelagic fish with densities of at least one individual per meter^3. The technique relies upon acoustic waveguide propagation in the continental shelf. Here, trapped modes dominate propagation and suffer only cylindrical spreading loss rather than the spherical loss suffered in free-space transmission or short-range propagation in the ocean. In contrast, standard methods for fish surveyance involve line transect measurements from slow moving research vessels that significantly under-sample fish distributions in time and space, leaving an incomplete behavioral picture. The implications of this bioclutter phenomenon on the Navy's long range active sonar operations in continental shelf environments are discussed.

FRIDAY AFTERNOON, 28 MAY 2004 NEW YORK BALLROOM B, 1:25 TO 5:15 P.M.

Session 5pEA

Engineering Acoustics and Underwater Acoustics: Autonomous Underwater Vehicle Acoustics: Part 2: Hardware and Devices

Thomas R. Howarth, Cochair
NAVSEA Newport, 1176 Howell Street, Newport, Rhode Island 02841

Henrik Schmidt, Cochair
Department of Ocean Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Chair's Introduction—1:25

Invited Papers

1:30

5pEA1. Broadband, lower-frequency acoustic projector design for AUV applications. Thomas R. Howarth (U.S. Navy, Newport, RI and Washington, DC, thowarth@ccs.nrl.navy.mil), Kim C. Benjamin, Dehua Huang (Naval Sea Systems Command, Div. Newport, Newport, RI), James F. Tressler (Naval Res. Lab., Washington, DC), and Walter Carney (Naval Sea Systems Command, Crane, IN)

Autonomous underwater vehicles (AUVs) offer a difficult platform for mounting lower-frequency acoustic projectors. Researchers at various U.S. Navy laboratories have been coordinating their interests to address the concepts and reduction to practice of lower frequency and broader operating frequency bands transduction technologies. Recent efforts have resulted in the development of cymbal-based acoustic transmitting devices to operate over an extended frequency range of 1 to 25 kHz within a package that can be housed inside of a BlueFin AUV. This presentation will discuss the concepts, design, development, and integration of the cymbal-based modular panels and how they are being configured to meet these applications. *In situ* data, including in-water calibration of acoustic performance, will be provided. Discussions for ongoing work and future directions will conclude the presentation. [Work supported by the ONR.]

1:50

5pEA2. 1-3 Piezocomposite transducers for AUV applications. Brian Pazol, Ken Lannaman, and Barry Doust (Mater. Syst., Inc., 543 Great Rd., Littleton, MA 01460)

Sonar systems on board AUVs present interesting challenges to the transducer designer because of their small size, low weight requirements, and limited available power. 1-3 piezocomposite transducers offer many performance characteristics which make them ideal for deployment in AUVs. Piezocomposite transducers are light weight, have broad bandwidth, have high efficiency, and can be conformed to fit the curvature of the vehicle. The broad bandwidths and low sidelobes made possible by piezocomposites result in sharper images with less distortion. The piezocomposite material is mechanically robust and can survive the rigors of normal operations as well as AUV deployment and retrieval. In addition, the conformal configuration substantially reduces hydrodynamic drag. As a conformal array, there is nothing to get knocked off during deployment and retrieval operations, or entangled with natural or man-made objects suspended in the water column. This contributes directly to improving the operational endurance of the AUV system, thereby enhancing overall system utility. MSI has produced and tested a variety of piezocomposite transducers for use in obstacle avoidance, mine hunting, and acoustic communications. An overview of piezocomposites and recent results of piezocomposite transducers will be presented.

2:10

5pEA3. Acoustic communications and autonomous underwater vehicles. Lee Freitag, Matthew Grund, James Preisig (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, lfreitag@whoi.edu), and Milica Stojanovic (MIT, Cambridge, MA 02139)

Acoustic communications systems used on autonomous underwater vehicles (AUVs) provide supervisory control, access to real-time data, and also allow multiple vehicles to cooperate in undertaking adaptive sampling missions. However, the use of acoustic systems on AUVs presents special challenges because of limited space for optimal placement of transducers, and potential conflicts with other acoustic systems such as side-scan sonars and transponders. In addition, radiated and structure-borne acoustic interference from thrusters and actuators reduces the sensitivity of on-board receivers. Recent work in acoustic communications and AUVs has included combining some navigation functions into communications equipment, development of operating modes that remove conflicts between different subsystems, design of vehicle components to avoid or remove interference, and other approaches to improving performance. While these efforts have been successful for specific installations, many challenges remain. This talk addresses problems and solutions for supervised and completely autonomous multi-vehicle communications to support complex AUV missions. Also presented are recent results which demonstrate that acoustic communications can be used successfully on a variety of AUV platforms for many different applications. [Work supported by ONR.]

2:30

5pEA4. Acoustic pressure-vector sensor array. Dehua Huang, Roy C. Elswick (NUWC, Newport, RI 02841), and James F. McEachern (ONR, Arlington, VA 22217)

Pressure-vector sensors measure both scalar and vector components of the acoustic field. December 2003 measurements at the NUWC Seneca Lake test facility verify previous observations that acoustic ambient noise spectrum levels measured by acoustic intensity sensors are reduced relative to either acoustic pressure or acoustic vector sensor spectrum levels. The Seneca measurements indicate a reduction by as much as 15 dB at the upper measurement frequency of 2500 Hz. A nonlinear array synthesis theory for pressure-vector sensors will be introduced that allows smaller apertures to achieve narrow beams. The significantly reduced ambient noise of individual pressure-vector elements observed in the ocean by others, and now at Seneca Lake, should allow a nonlinearly combined array to detect significantly lower levels than has been observed in previous multiplicative processing of pressure sensors alone. Nonlinear array synthesis of pressure-vector sensors differs from conventional super-directive algorithms that linearly combine pressure elements with positive and negative weights, thereby reducing the sensitivity of conventional super-directive arrays. The much smaller aperture of acoustic pressure-vector sensor arrays will be attractive for acoustic systems on underwater vehicles, as well as for other applications that require narrow beam acoustic receivers. [The authors gratefully acknowledge the support of ONR and NUWC.]

2:50

5pEA5. Advanced sonar array concepts for small underwater vehicle applications. Kim C. Benjamin (NAVSEA Div. Newport, 1176 Howell St., Newport, RI 02841)

Low-frequency acoustic reception and high directivity are difficult to obtain simultaneously using small undersea (robot) platforms. The dimensions and geometry of the hydrodynamic shapes that are required for efficient transit through the water preclude the use of broad apertures required for low-frequency directional applications. This paper will present array concepts that could be considered the next step in developing low-frequency directional acoustic payloads for undersea research and exploration. The talk will include results of previous Navy experimental demonstrations of wide aperture arrays for small vehicles as well as future advanced sonar packaging concepts that rely on new transduction technologies. [Work supported by the U.S. Navy.]

3:10–3:30 Break

Contributed Papers

3:30

5pEA6. Compact, high power, energy efficient transmit systems for UUVs using single crystal transducers. Harold Robinson (NUWC Div. Newport, Code 2132, Bldg. 1170, 1176 Howell St., Newport, RI)

UUVs are currently being designed to perform a multiplitude of tasks in ocean exploration and Naval warfighting. Many of these tasks require the use of active acoustic projectors, and many may require the UUV to operate independently for hours, days, or even weeks. In order for a UUV to be as versatile as possible, its active transmit system must be versatile as well, implying that broad acoustic bandwidths are a must. However, due to size and battery life limitations, this broadband system must also be compact and energy efficient. By virtue of their extraordinary material properties, ferroelectric single crystals are the ideal transduction material for developing such broadband systems. The effect of their high coupling factor on transmit systems shall be illustrated by showing the dramatic impact on amplifier size, power factor, and acoustic response that is possible using these materials. In particular, a transducer built with these materials can be well matched to the power amplifier, i.e., 80% or more of the amplifier power reaches the transducer, over decades of frequency. Measured results from several prototype single crystal transducers shall be presented to demonstrate that the theoretical gains are actually realizable in practical devices. [Work sponsored by DARPA.]

3:45

5pEA7. A dual-piston ring-driven X-spring transducer. Alexander L. Butler, John L. Butler (Image Acoust., Inc., 97 Elm St., Cohasset, MA 02025, abutler@imageacoustics.com), Robert L. Pendleton, and Richard M. Ead (Naval Undersea Warfare Ctr., Newport, RI 02841)

Tonpilz transducers generally consist of a stack of piezoelectric material sandwiched between a single piston and an inertial tail mass or between two pistons. The result is a transducer with a large length-to-diameter ratio. The X-spring transducer design, based on U.S. Patent 4 845 688, allows a means for a shorter transducer length through an orthogonal piezoelectric drive system coupled to the pistons by lever arms. We present here a low-frequency, dual-piston piezoelectric ceramic ring driven version with a length of only 10 in. and a diameter of 19 in. Both single-element and two-element array results are presented. The measured response is shown to be in agreement with the finite-element model with a

smooth, wideband 300- to 550-Hz response for this dual-piston, ring-driven X-spring transducer. [Work supported by a Phase II SBIR, through NUWC, Newport, RI 02841.]

4:00

5pEA8. Acoustic cymbal performance under hydrostatic pressure. Kirk E. Jenne, Dehua Huang, and Thomas R. Howarth (NAVSEA Div. Newport, 1176 Howell St., Newport, RI 02841-1708, JenneKE@npt.nuwc.navy.mil)

Continual awareness about the need to develop light-weight, low-volume, broadband, underwater acoustic projector and receive arrays that perform consistently in diverse environments is evident in recent Navy acoustic system initiatives. Acoustic cymbals, so named for resemblance to the percussive musical instruments, are miniature flextensional transducers that may perhaps meet the performance criteria for consistent performance under hydrostatic pressure after modifications in the design. These acoustic cymbals consist of a piezoceramic disk (or ring) bonded to two opposing cymbal-shaped metal shells. Operating as mechanical transformers, the two metal shells convert the large generative force inherently within the disk's radial mode into increased volume displacement at the metal shell surface to obtain volume displacement that translates into usable source levels and/or sensitivities at sonar frequencies in a relatively broad band. The air-backed design for standard acoustic cymbal transducers presents a barrier to deepwater applications. A new acoustic cymbal design for high-pressure applications will be presented for the first time. This practical pressure compensation is designed to diminish the effects of hydrostatic pressure to maintain consistent acoustic cymbal performance. Transmit and receive performance data, determined at the Naval Undersea Warfare Center's (NUWC) Acoustic Pressure Tank Facility (APTF), is presented.

4:15

5pEA9. An acoustic array to be towed behind an autonomous underwater vehicle. Jason D. Holmes, William M. Carey (Boston Univ., Boston, MA 02215, jholmes@bu.edu), and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The use of a towed array from a small autonomous underwater vehicle, AUV, such as the WHOI-Remus vehicle is discussed as a valuable ocean acoustics measurement tool for of 3-D characterization of shallow water regions. The feasibility of towing a 10-m-long, small-diameter fluid-filled hydrophone array behind Remus was investigated with a laboratory-prototype array and preliminary tests from the WHOI pier. Preliminary results on array self-noise are presented for low tow speeds. The flow noise, vehicle noise, and other unwanted signal degrading noise sources are examined, discussed, and shown by theoretical arguments to be reduced to low levels by proper array design. This paper examines current technology for such an array with digital sampling and recording equipment incorporated in the Remus vehicle. Consideration is also given to the deployment problem of array on an autonomous vehicle. [Work supported by the College of Engineering, Boston University.]

4:30

5pEA10. A framework of concurrent navigation and seabottom targets detection using acoustic sensors on AUVs. Te-Chih Liu and Henrik Schmidt (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

The use of Autonomous Underwater Vehicles (AUVs) for Mine Counter Measures (MCMs) is an area of active recent research. The excellent mobility of AUVs allows for multi-aspect sonar view of the targets for improved detection, tracking, and classification. However, the uncertainty of the platforms and associated target localization degrade the detection ability by AUVs. Furthermore, the weak signals from buried targets is another severe problem making detection by a single measurement impossible. A new acoustic sensing framework is proposed for detection and tracking seabottom targets based on the track-before-detect (TBD) technique. In contrast to the traditional methods, TBD tracks possible targets before the detection is declared. There are several advantages by using this framework: (1) Compared to the traditional method of detection followed by tracking, higher detection probability is achieved for dim target detection due to the integrated detection metric of TBD. (2) Compared to the multiple hypothesis tracker (MHT) method, instead of searching a diverge hypotheses tree, TBD speeds up the searching and decreases the computational load, which makes onboard implementation feasible. (3) The stochastic models of uncertainties of targets and AUVs are based on the Bayesian framework, and thus, it is easy to apply various recursive estimators such as the Kalman filter or particle filter for tracking individual targets as well as the AUV platforms. Results of a successful application of this method in the GOATS2002 experiment are demonstrated. [Work supported by ONR and NATO Undersea Research Centre.]

4:45

5pEA11. Acoustic navigation for autonomous underwater vehicles. Wen-Bin Yang and T. C. Yang (Naval Res. Lab., Code 7120, 4555 Overlook Ave. SW, Washington, DC 20375, wyang@wave.nrl.navy.mil)

To effectively use groups of autonomous underwater vehicles (AUVs), an accurate navigation capability must be developed. We describe an acoustic navigation technique that uses an acoustic communications probe signal to estimate the range, range rate and direction of arrival signal. The acoustic signal may be transmitted from one vehicle or node (e.g., AUV, ship or fixed mooring) to another. The technique was tested during the RAGS03 experiment. A towed acoustic communications receiver that replicated the motion of an AUV and a fixed source node were used. The receiver was navigated with a global position system (GPS). The range, range rate and direction of arrival signal have been estimated using the acoustic signal received on a horizontal line array. The results are compared to the known GPS source and receiver positions. The navigation technique will be discussed. Its advantages over the traditional navigation by multiple transponders will be outlined. [Work supported by the ONR.]

5:00

5pEA12. Experimental modeling of the acoustic signature of an AUV. Joseph M. Cuschieri (Lockheed Martin Corp., MS2 Undersea Systems, Riviera Beach, FL 33404)

In this presentation test tank results from an experimental analysis of the acoustic signature of an Ocean Explorer class AUV are presented. The results are from measurements performed in a reverberant test tank on an AUV model, and of an AUV under typical operating conditions. The main source of excitation is the propulsion module or "podule." Inside the podule are the propulsion motor and the motors for the control surfaces with penetrations for the main propulsion shaft and the shafts for the control surfaces. Different operating conditions and different mountings for the main propulsion and control module (podule) inside the AUV are considered. The influence of the propeller, the mounting of the podule, and covering the podule with a compliant layer are considered in the measurements. It is shown that for the type of AUV considered here, the mounting of the podule is not very significant and that significant energy is transferred through the water trapped between the podule and the AUV hull. Furthermore, the propeller has a significant influence on the acoustic signature. [Work supported by ONR.]

Session 5pNS

Noise: Noise in Large Cities II

Daniel R. Raichel, Chair
2727 Moore Lane, Fort Collins, Colorado 80526

Chair's Introduction—1:15

Invited Papers

1:20

5pNS1. The effect of rail sources on noise-sensitive construction in Chicago. Brian L. Homans (Shiner + Assoc., Inc., One N. Franklin, Ste. 2025, Chicago, IL 60606, bhomans@shineracoustics.com)

Because of Chicago's dominance as a rail center in the 19th and 20th centuries and the miles of track present, parcels of land available for residential and other noise-sensitive construction frequently lie adjacent to freight, commuter, and Chicago Transit Authority (CTA) tracks. As A-weighted sound levels commonly exceed 105 dBA underneath CTA elevated tracks, special methods of noise abatement for high-rise and other construction to address high noise levels are necessary. Also discussed will be methods of establishing acoustical criteria appropriate for commercial and residential interior spaces.

1:40

5pNS2. 24 hours of noise in a large city, problems and solutions. Leslie Blomberg (Noise Pollution Clearinghouse, P.O. Box 1137, Montpelier, VT 05601)

This paper begins by comparing a contemporary urban noise study to studies done in the 1970s. Several changes in the urban soundscape are noted, as well as areas in which progress has and has not been made in quieting urban soundscapes. Finally, the challenges to quieting urban areas are presented and potential solutions are identified.

2:00

5pNS3. Archaeological acoustics—A guide to trends in community noise levels. Daniel R. Raichel (Eilar & Assoc., 2727 Moore Ln., Fort Collins, CO 80526, draichel@eilarassociates.com), Bennett M. Brooks (Brooks Acoust. Corp., Vernon, CT 06066), and David Lubman (David Lubman & Assoc., Westminster, CA 92683)

Archaeological acoustics may be defined as assessment of the acoustical situations of the past by the use of scientific methods in examining the available remaining evidence. These principles may be applied to the study of historic trends in community noise by measuring the types of noise sources prevalent at the time, estimating the noise level of the environment on the basis of information as to types and number of noise sources, amount of traffic and industry, types of road surfaces, and layout of surrounding buildings and local landscaping. For instance, the clopping of horses and wagons passing by can be measured with modern instruments at current-day rodeos, parades, and exhibitions. The noise levels can be transposed to a bygone scene through acoustical analyses, possibly with computer-aided methodologies. Archives can be consulted for population density, details of buildings that existed at the time, paving types of streets, presence of sidewalks, amount of commercial and pedestrian traffic, etc. Estimates of the noise levels in major cities for specific periods of time can lead to detection of noise trends over a period of centuries. Several noise trend estimate examples are illustrated.

2:20

5pNS4. Evaluation of the noise pollution in urban parks of Curitiba, Brazil. Andressa C. Ferreira, Fabiano B. Diniz, Elaine C. Paz, and Paulo T. Zannin (Environ. Acoust. Lab., Dept. of Mech. Eng., UFPR, Brazil)

This work shows a study about the noise pollution found in six urban parks of Curitiba, Paran, Brazil. The equivalent noise levels (Leq) have been measured in points spread throughout the park, and interviews have been conducted with some park visitors. It has been found out that 52.48% out of the measurement sites did not satisfy the Municipal Law no. 10,625, which states the noise emission level of 55 dB(A) as the limit value for green areas. The results of the questionnaires applied to the local visitors have showed that 39% out of the interviewed people used to visit the park every day and that 75% out of them seek for the realization of a physical activity. During the realization of their activities in the parks, 22% out of the interviewed people pointed to the noise pollution as the source of annoyance and 28% out of them pointed the local security. In this sense, it has been verified that half of the analyzed parks were inserted in acoustically polluted areas, which incurs a real state depreciation in their vicinities.

2:40–2:45 Break

2:45

5pNS5. Public acceptance of urban rotorcraft operations. Michael A. Marcolini, Clemans A. Powell, and Joe W. Posey (NASA Langley Res. Ctr., Hampton, VA 23681)

Even though tiltrotor operations from city center to city center could greatly shorten travel times over moderate ranges, public opposition to intense urban rotorcraft activity has kept this possibility from being realized. One significant factor in this opposition is rotorcraft noise. Over the last 25 years, NASA has explored the subjective response to rotorcraft noise and developed low noise design concepts and noise abatement flight procedures. While low noise designs can be applied for future rotorcraft, this is not an effective near-term means of reducing rotorcraft noise, because of the costs associated with replacement of helicopter rotor blades. Recent noise abatement research, which has been focusing on the development of tools and techniques to facilitate the design of quieter flight procedures for existing vehicles, has much more immediate application. While very little subjective response work has occurred recently, prior work at NASA in this area from the 1970s and 1980s is discussed. Lastly, thoughts on future research areas that might help improve the public acceptance of rotorcraft will be described.

3:00

5pNS6. Study of the daily environmental urban noise levels in San Juan, Puerto Rico. José A. Alicea-Pou, Olga Viñas-Curiel, Wanda Cruz-Vizcarrondo, Daniel Hernandez-Dávila (Noise Control Area, Environ. Quality Board of Puerto Rico), and Jorge Rocaford (School of Architecture, Univ. of Puerto Rico)

The study of the urban environmental noise levels in the city of San Juan is one of the first steps into understanding the noise behavior of the largest and denser populated zones in Puerto Rico. For the year 2002 San Juan had 36% of all citizen noise complaints reported for the entire island. The aim of the study was to monitor for 24 continuous hours the levels of urban sounds and noise for different locations in the city and describe the general tendencies regarding those levels for residential, commercial, recreational or industrial areas. The sampling was done with four (4) units of the Norsonic NOR-121 sound analyzer that were installed temporarily for 24 hours in multiple locations chosen randomly around the city. The study is expected to continue during the next 2 years. The preliminary data showed that the noise levels in residential areas with a high traffic routine fluctuated in levels that were consistently high all day and most of the night, dropping around 1:00 a.m. and rising again around 5:00 a.m. The results of this and other projects would be used to build and implement a National Noise Plan for Puerto Rico for 2005–2007.

3:15

5pNS7. Study on the correlation between the old city and the newly developed city regarding audible sound types and sound levels. Takeshi Tokashiki (Faculty of Eng., Univ. of the Ryukyus, 1 Senbaru Nishihara cho Nakagami gun Okinawa PRef. Japan, tkashiki@tec.u-ryukyu.ac.jp), Yasuhiro Yamashita, and Naoki Takagi (Shinshu Univ., Nagano, Japan)

Sound-level data are the most important among the appraisal data of environmental noise. However, to appraise diversified environment like that of the present times, a more detailed appraisal on environmental noise could be obtained by conducting subjective appraisal in addition to the physical noise level. In this research, we have surveyed subjective appraisal, audible sound types and sound level (LA_{eq}) along boulevard, commercial, and residential areas in the old city and the newly developed city. The subjective appraisal by examinees shows that of the 100% of noise along boulevards, traffic sound occupies 80%; of the 60% of noise in commercial areas the traffic sound occupies 65%; and of the 40% of noise in residential areas the traffic sound occupies 50%. We have studied the correlation between the result of the same surveyed in the old city and the newly developed city. There have been fewer study cases about the correlation between sound types and subjective appraisal; therefore, in this paper we have studied this correlation.

3:30

5pNS8. Comparative analysis of the urban noise between two different areas in the city of Curitiba, PR. Elaine C. Paz, Andressa C. Ferreira, and Paulo T. Zannin (Environ. Acoust. Lab., Dept. of Mech. Eng., UFPR, 81531-990, Curitiba, PR, Brazil)

The purpose of this work is to analyze the urban noise perception comparatively in the inhabitants of a residential area (neighborhood) and a mixed area (center), in the city of Curitiba, PR, in order to characterize two different situations: (1) acoustically ideal urban environment; and (2) acoustically polluted urban environment. For that, subjective and objective evaluations were accomplished, where an aleatory sample of each area was submitted to a survey. In the objective evaluation, the medium equivalent sound levels calculated were 53.50 dB(A) and 72.90 dB(A) for the neighborhood and center, respectively. The parameters used for comparison of the calculated medium equivalent sound levels where the values of 55.00 dB(A) (Municipal Law No. 10.625) and 65.00 dB(A) (WHO), in the period of the day for residential areas. The interpretation of the subjective results verified that the central zone inhabitants have an annoyance perception bigger than the residential zone inhabitants. The interpretation of the objective results classified the neighborhood and center areas as acoustically control zone and acoustically polluted zone, respectively, according to the adopted parameters. Starting from the comparison between these two areas, it was defined that both can be classified as reference factor for other evaluations.

Session 5pPA

Physical Acoustics: Wave Propagation: Inhomogeneous Media

Emmanuel Bossy, Chair

Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Contributed Papers

1:00

5pPA1. Ultrasonic inhomogeneous waves: Three decades of fascination. Nico F. Declercq (Soete Lab., Ghent Univ., Sint Pietersnieuwstraat 41, B-9000 Ghent, Belgium, nicof.declercq@ugent.be), Rudy Briers (KATHO, Torhout, Belgium), Joris Degrieck (Ghent Univ., B-9000 Ghent, Belgium), and Oswald Leroy (IRC-KULAK, Krotrijk, Belgium)

Inhomogeneous waves are mathematically described as pure homogeneous plane waves except for the important fact that all wave parameters, i.e., wave vector, polarization vector and frequency, can be complex valued. Even though the existence of a complex wave vector was known for a long time, Henry Cooper was the first to study such waves and their interaction with interfaces. It was not until 1980 that researchers became aware of the fact that inhomogeneous waves are mandatory phenomena in the description of reflection/transmission phenomena between lossy media. Later, it was shown that inhomogeneous waves also formed natural building blocks of bounded beams and formed a physical explanation for the Schoch effect. This paper describes this development from a historical point of view and describes scientific properties (e.g., their polarization, dispersion) whenever needed in order to understand the evolution from 1980 until the present. All topics of inhomogeneous wave research are taken into account, such as waves in viscoelastic solids and liquids, thermo-viscous liquids and solids, anisotropic viscoelastic materials, periodically rough materials, the features of complex frequency and the experimental generation of inhomogeneous waves. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.).]

1:15

5pPA2. Acoustic pulse propagation through stable atmospheric boundary layer: Theory and experiment. Igor Chunchuzov, Sergey Kulichkov, Alexander Otrezov, and Vitaly Perepelkin (Obukhov Inst. of Atmospheric Phys., 3 Pyzhevskii, Moscow, oksana@achilles.net)

Mesoscale wind speed and temperature fluctuations with periods from 1 min to a few hours significantly affect a variability and turbulent regime of stable atmospheric boundary layer (ABL). Their statistical characteristics are still poorly understood, although the knowledge of such statistics is required when modeling sound propagation through stable ABL. Several field experiments have been conducted to study the influence of mesoscale wind speed fluctuations on acoustic pulse propagation in stable ABL. The results of these experiments are presented in this work. A special acoustic source was used to generate acoustic pulses due to a detonation of air-propane mixture with a repetition period of 1 min or 30 s. The mean wind speed profiles and mesoscale wind fluctuations were measured by Doppler sodar up to a height of 300 m, and by anemometers placed on a 56-mast. From the measurements of the pulse travel time fluctuations at different distances from the source the statistical characteristics of the mesoscale wind fluctuations such as frequency spectra, coherences, horizontal phase speeds and scales have been obtained. Some of the obtained results are interpreted with a recently developed model of internal wave spectrum in a stably stratified atmosphere.

1:30

5pPA3. Time-domain equations for sound propagation in or reflection from rigid porous media. Vladimir E. Ostashev (NOAA/Environ. Technol. Lab., 325 Broadway, Boulder, CO 80305), D. Keith Wilson (U.S. Army Engineer Res. and Development Ctr., Hanover, NH 03755), and Sandra L. Collier (U.S. Army Res. Lab., Adelphi, MD 20783)

In acoustics, many absorbing materials can be considered as porous media with rigid frames. Therefore, studies of sound propagation in or reflection from rigid porous media are important in many applications. These studies are usually done in the frequency domain where equations describing sound propagation in or reflection from porous media are well established. However, there are also problems where such equations are needed in the time domain. Examples include studies of material properties by acoustic impulses and finite-difference time-domain (FDTD) simulation of sound propagation in the presence of absorbing surfaces. In this paper, using Wilson's relaxation model for acoustical properties of porous media [D. K. Wilson, Appl. Acoust. **50**, 171–188 (1997)], we derive a closed set of time-domain, integro-differential equations for the sound pressure and particle velocity in rigid porous media. In the limiting cases of low and high frequencies, these equations coincide with those known in the literature. Furthermore, in these limiting cases, using the relation model, we derive time-domain boundary conditions for sound reflection from a porous medium. This result can be used in FDTD simulation of sound propagation outdoor. [Work supported by the DoD High Performance Computing Modernization Office and ERDC-CRREL.]

1:45

5pPA4. Frequency dependent P-wave speed in a porous medium with aligned fractures. Miroslav Brajanovski, Boris Gurevitch (Dept. of Exploration Geophys., Curtin Univ. of Technol., G.P.O. Box U1987, Perth WA 6845, Australia), and Michael Schoenberg (CSIRO Petroleum, Bentley WA 6102, Australia)

In an elastic medium, fractures can be modeled as thin layers, the elastic stiffnesses of which approach zero as the volume fraction of the fractures $h_f \to 0$. This yields linear slip theory [M. Schoenberg, J. Acoust. Soc. Am. **68**, 1516–1521 (1980)], shown to be a robust way to account for the acoustic effect of fracturing. From Norris' [J. Acoust. Soc. Am. **94**, 359–370 (1993)] dispersion relation for alternating porous layers, fractures must be modeled similarly in porous media. However, another fracture parameter of great interest is permeability. If fracture permeability is taken to be $\mathcal{O}(h_f^{-1})$ or taken to be independent of h_f, one arrives at a dispersion relation for the fast P-wave dependent on porous background properties and a real excess compliance which takes the fractures into account. However, if fracture permeability is assumed to be small, and is taken to be $\mathcal{O}(h_f)$, the P-wave dispersion dependence on background parameters remains the same, but the term accounting for the fractures is frequency dependent. Only in the zero frequency limit does this result agree with that of the other two cases.

2:00

5pPA5. Acoustical properties of gravel. Keith Attenborough and Olga Umnova (Dept. of Eng., The Univ. of Hull, Cottingham Rd., Hull HU6 7RX, UK)

Gravel is an example of a rigid-porous granular material. It is available with several mean stone sizes and hence with a range of flow resistivity. The flow resistivity of gravel varies significantly with flow velocity. Data for low- and high-amplitude impedance are presented and compared with predictions of linear and nonlinear theories based on the Johnson-Allard model for rigid-porous media. Comparisons are made also between data and predictions for shock wave reflection and transmission at single- and multiple-layer gravel surfaces including recent data obtained with laser-generated acoustic shocks. Tolerable agreement is obtained between data and predictions. It is found that a low flow resistivity gravel layer has a reflection coefficient that has a minimum as the incident pressure is increased. A layered system that offers the lowest reflection coefficient at linear sound pressures does not continue to do so as the incident pressure is increased. [Work supported by USACE ERDC BT25 program.]

2:15

5pPA6. Is a moving average field a proper macro scale measure? John J. McCoy (The Catholic Univ. of America, Washington, DC 20064)

A number of derivations of the Biot equations governing the acoustics of fluid-filled, porous solids, which are based on a more complete formulation that applies to all length scales, accept moving averages of the response fields that enter the more complete formulation as the response fields that enter the Biot equations. This raises a question: Does a moving average field incorporate *only* macro scale variation? A moving average field is presented as one extreme of a class of fields that is formed from sets of discrete local spatial averages. The set of local spatial averages for a moving average field is accomplished for locations that are separated by a vanishing distance. The opposite extreme of a field of local averages is formed from a set of local averages accomplished for locations that are separated by a distance equal to a linear measure of the region of the spatial average. Explicitly demonstrated is that the moving average of a field that contains both macro and micro scale variation *will itself contain* both macro and micro scale variation. The relative suppression of the micro scale variation compared to the macro scale variation, which obtains in one representation of the moving average field, is only apparent; the micro scale variation can be recovered by an appropriate signal processing. This is in contrast to a wavelet-defined field of local averages, an example of the other extreme, for which the suppression of the micro scale variation is absolute. The issue is significant for a derivation of a prediction model that purports to output macro scale response fields.

2:30

5pPA7. Doublet mechanics in acoustics. Christopher N. Layman, Jr. and Junru Wu (Dept. of Phys., Univ. of Vermont, Cook Physical Sci. Bldg., 82 University Pl., Burlington, VT 05405-0125, jwu@zoo.uvm.edu)

Doublet mechanics (DM) presents a micromechanical approach to materials, whose properties allow for a granular characterization. In DM a pair (doublet) of interacting granules represents the fundamental unit of the material; these two granules are either in contact or connected via a thin adhesive layer. The actions of a doublet are described via a convergent Taylor series, which can be truncated at any level dependent upon how discrete the system is. This is the flexibility of DM, meaning it allows for the use of a single framework to describe materials from low discreteness (i.e., animal tissue) to high discreteness (i.e., solids). Some preliminary work [J. Liu and M. Ferrari, Dis. Markers **18**, 175–183 (2002); J. Wu, C. Layman, and J. Liu, J. Acoust. Soc. Am. (in press)] has produced encouraging results in the application of DM to the ultrasonic characterization of tissue.

2:45

5pPA8. Computation of the dynamic thermal properties of a three-dimensional unit cell of porous media by Brownian motion simulation. Camille Perrot (ENTPE and GAUS), Xavier Olny (ENTPE, DGCB URA CNRS 1652, Vaulx-en-Velin Cedex, France), Raymond Panneton, and Richard Bouchard (GAUS, Univ. de Sherbrooke, Canada)

Acoustic dissipation in porous media is mainly due to viscous and thermal mechanisms that occur in the pores of the microstructure. The purpose of this study is the determination of the macroscopic dynamic acoustic bulk modulus and thermal permeability of real foams from a local scale approach. To achieve this goal, two distinct steps are followed. First, the local geometry of a real foam is obtained using computed microtomography (μCT), then a periodic and regularly paving space tetrakaidecahedron cell is identified from the microstructure. Second, the heat equation is solved for the geometrical model. The paper provides a three-dimensional application of the efficient simulation technique of Brownian motion proposed by Torquato *et al.* for steady state diffusion-controlled problems [Appl. Phys. Lett. **55**, 1847–1849 (1989)] and adapted by Lafarge [Poromechanics **II**, 708 (2002)] in a bi-dimensional case. The influence of the model's microstructural details (anisotropy, and struts junction and cross-section) on the macroscopic properties are studied. The predictions of the macroscopic properties using this local scale approach are then compared to experimental measurements.

3:00

5pPA9. Generation of subharmonic and difference frequency acoustic waves in a water-saturated sandy sediment. Byoung-Nam Kim, Kang Il Lee, Suk Wang Yoon, and Bok Kyoung Choi (Dept. of Phys., Sung KyunKwan Univ., Suwon 440-746, Republic of Korea, swyoon@skku.ac.kr)

Generation of subharmonic and difference frequency acoustic waves in a water-saturated sandy sediment was investigated. Subharmonic frequency acoustic wave was observed to be generated due to the nonlinearity of water-saturated sandy sediment when the fundamental frequency acoustic wave exceeded a certain threshold pressure amplitude. Pressure spectrum level of the subharmonic frequency acoustic wave linearly increased as the driving acoustic pressure amplitude increased. The pressure level was 10 dB higher than the background noise level. Generation of the difference frequency acoustic wave was also observed by the collinear acoustic waves with two different fundamental frequencies. The pressure level of the difference frequency acoustic wave was 5 dB higher than the background noise level. It seems very useful to evaluate the nonlinear parameter of water-saturated sandy sediment without disturbing the sediment. Such nonlinear acoustic responses of water-saturated sandy sediment can be utilized for the diagnosis of marine gassy sandy sediment.

3:15–3:30 Break

3:30

5pPA10. Diffusion of high-frequency energy in fluid-saturated porous media. Eric Savin (Structural Dynam. and Coupled Systems Dept., ONERA, 29 Ave. de la Div. Leclerc, 92322 Châtillon cedex, France)

The modern mathematical theory of microlocal analysis shows that the energy associated with the high-frequency solutions of hyperbolic partial differential equations (such as the wave or the Navier equations) satisfy Liouville-type transport equations, or radiative transfer equations for randomly heterogeneous media. For long propagation times the latter can be approached by diffusion equations. Some classical results of the structural acoustics literature about the heat conduction analogy and the statistical energy analysis of structural dynamics at higher frequencies are recovered in this process. The purpose of this communication is to focus on such a diffusive regime for isotropic, fluid-saturated porous media. More specifically, we have derived the diffusion parameters (transport mean-free path and diffusion constant) for such media. Our model considers Biot's equations of poroelasticity, where thermal and viscous effects are modelized by dynamic tortuosity and compressibility with singular memory kernels. The macroscopic bulk modulus and density of the dry solid phase are assumed to be homogeneous random processes, while tortuosity and porosity remain constant.

3:45

5pPA11. Surface and plate mode stimulation in piezoelectric materials in the framework of inhomogeneous wave theory. Nico F. Declercq, Joris Degrieck (Soete Lab., Ghent Univ., Sint Pietersnieuwstraat 41, B-9000 Ghent, Belgium, nicof.declercq@ugent.be), and Oswald Leroy (IRC-KULAK, Kortrijk, Belgium)

Inhomogeneous waves are already a well established concept in ultrasonics. Whenever damping is present, such waves become inevitable when reflection and transmission phenomena have to be described. The concept of inhomogeneous waves has been established in the study of acoustics in liquids, isotropic solids and also anisotropic solids. It has been shown before that inhomogeneous waves are much better in stimulating surface waves and plate modes than plane waves and general bounded beams. Nevertheless the concept has not been introduced yet in the field of piezoelectric media. The current paper describes the interaction of inhomogeneous waves with piezoelectric solids of any kind of anisotropy and shows how surface and plate modes can be excited by means of such waves. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.).]

4:00

5pPA12. Diffraction of plane waves having complex frequency and the excitation of transient leaky Rayleigh waves. Nico F. Declercq, Joris Degrieck (Soete Lab., Ghent Univ., Sint Pietersnieuwstraat 41, B-9000 Ghent, Belgium, nicof.declercq@ugent.be), and Oswald Leroy (IRC-KULAK, Kortrijk, Belgium)

Harmonic plane waves are not advantageous for stimulation of leaky Rayleigh waves whereas bounded beams are better suited. Furthermore, past studies have revealed that it is even better to apply incident harmonic inhomogeneous waves, both on smooth and on rough surfaces. If the right inhomogeneity is used, very strong excitation is to be expected. However, inhomogeneous waves are difficult to generate and are far from practical outside the laboratory boundaries. The current paper describes how incident harmonic plane waves having a complex frequency are also excellent tools for stimulating leaky Rayleigh waves, if the technique is combined with the use of a periodically rough surface. Such transient plane waves are much more practical to generate experimentally whence the usability of the technique is more realistic compared with the classical use of harmonic inhomogeneous waves. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.).]

4:15

5pPA13. Nonlinear modal method of crack localization. Lev Ostrovsky (Zel Technologies/NOAA ETL, Boulder, CO), Alexander Sutin (Stevens Inst. of Technol., Hoboken, NJ), and Andrey Lebedev (Inst. of Appl. Phys., Nizhniy Novgorod, Russia)

A simple scheme for crack localization is discussed that is relevant to nonlinear modal tomography based on the cross-modulation of two signals at different frequencies. The scheme is illustrated by a theoretical model, in which a thin plate or bar with a single crack is excited by a strong low-frequency wave and a high-frequency probing wave (ultrasound). The crack is assumed to be small relative to all wavelengths. Nonlinear scattering from the crack is studied using a general matrix approach as well as simplified models allowing one to find the nonlinear part of crack volume variations under the given stress and then the combinational wave components in the tested material. The nonlinear response strongly depends on the crack position with respect to the peaks or nodes of the corresponding interacting signals which can be used for determination of the crack position. Juxtaposing various resonant modes interacting at the crack it is possible to retrieve both crack location and orientation. Some aspects of inverse problem solutions are also discussed, and preliminary experimental results are presented.

4:30

5pPA14. Omnidirectional elastic bandgap in finite one-dimensional phononic systems. Betsabe Manzanares-Martínez (CIFUS, Univ. of Sonora, Hermosillo, Sonora 83190, Mexico, mbmm@cajeme.cifus.uson.mx), José Sánchez-Dehesa, Andreas Håkansson (Polytechnic Univ. of Valencia, E-46022 Valencia, Spain), Francisco Cervera (Polytechnic Univ. of Valencia, E-46022 Valencia, Spain), and Felipe Ramos-Mendieta (Univ. of Sonora, Sonora 83190, Mexico)

We have demonstrated experimentally the occurrence of omnidirectional elastic bandgaps in layered periodic structures. To date, omnidirectional reflection is a well-known phenomenon for the case of electromagnetic waves in photonic crystals [Y. Fink *et al.*, Science **282**, 1679 (1998)]. Theoretically a similar behavior of the elastic waves in phononic systems is predictable [D. Bria and B. Djafari-Rouhani, Phys. Rev. E **66**, 056609 (2002)]. For the experiment we used finite samples of alternating layers of Pb and epoxy. The binding medium is nylon, which is a material of high elastic velocity, a favorable condition to find the effect. The thicknesses of the layers were chosen in order to have omnidirectional gap at a few hundreds of kHz. We found good agreement between the experimental results and the theoretical treatment of the transmittance through finite samples and the phononic band structure. The effect is shown only for longitudinal impinging waves. [Work supported by CICyT of Spain and CONACyT of Mexico.]

4:45

5pPA15. Criteria for quasi-shear wavefront triplication in a tranversely isotropic material. Michael A. Schoenberg and Thomas M. Daley (Earth Sci. Div., Lawrence Berkeley Lab, 1 Cyclotron Rd., Berkeley, CA 94720)

Quasi-shear (qSV) wavefront triplication, classically known as shear wave bi-refringence, occurs in TI media when anellipticity parameter $E^2 \equiv BC - A^2$ differs significantly from 0 (equality is the elliptic case), where $A \equiv c_{13} + c_{55}$, $B \equiv c_{11} - c_{55}$, $C \equiv c_3 - c_{55}$ [Helbig and Schoenberg, J. Acoust. Soc. Am. **81**, 1235–1245 (1987)]. The region of the slowness curve corresponding to the triplicating region of the wavefront curve must be concave. Most common hexagonal crystals and all TI media long wavelength equivalent to isotropic layering exhibit positive anellipticity. In general, the exact condition for positive anellipticity triplication (the triplicating region then is between the 3-axis and its normal) requires the solution of a cubic equation [Peyton, *Elastic Wave Propagation in Transversely Isotropic Media* (Martinus Nijhoff, 1983)]; however, a good approximation for triplication is that $E^2 > KBCc_{55}/(c_{11} + c_{33})$, where K is almost always within a few percent of 1.39. For negative anellipticity media, a simpler case, triplication centers about the 3-direction when $A^2 > c_{11}C$ and/or about the normal to the 3-axis when $A^2 > c_{33}B$.

5:00

5pPA16. Generalized paraxial ray tracing derived from Riemannian geometry. David R. Bergman (Dept. of Phys., Saint Peter's College, 2641 Kennedy Blvd., Jersey City, NJ 07306, dbergman@spc.edu)

In 1973 R. White demonstrated that acoustic rays in a generic environment could be identified with the null geodesics of a pseudo-Riemannian manifold. A general set of paraxial (dynamic) ray equations, suitable for three dimensional ray tracing in a generic environment, is derived from the geodesic deviation equation used in general relativity and the geometric transmission loss of a ray bundle modeled from this equation. (When fluid motion is removed the paraxial ray procedure used in seismology emerges from the formalism.) The results are applied to time independent layered media where it is found that the standard ray integrals used in underwater acoustics as well as a generalized version of Snell's

law, originally derived by Kornhauser, emerge naturally as conserved quantities related to symmetries of the metric. Finally, when the results are applied to torsion free rays the deviation equation reduces to a scalar equation and the sectional curvature reduces to a simple expression depending on the derivatives of the sound speed and fluid velocity as well as the ray parameters allowing one to determine ray divergence or the development of focal points simply by checking the sign of a single term in the equation.

FRIDAY AFTERNOON, 28 MAY 2004 IMPERIAL BALLROOM B, 1:00 TO 5:00 P.M.

Session 5pSC

Speech Communication: Poster Session V

Nassima Abdelli-Beruh, Chair
Department of Speech-Language Pathology and Audiology, New York University, 719 Broadway, New York, New York 10003

Contributed Papers

All posters will be on display from 1:00 to 5:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:00 p.m. to 3:00 p.m. and contributors of even-numbered papers will be at their posters from 3:00 p.m. to 5:00 p.m.

5pSC1. Nonsensory factors in speech perception. Rachael F. Holt and Arlene E. Carney (Dept. of Commun. Disord., Univ. of Minnesota, 115 Shevlin Hall, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

The nature of developmental differences was examined in a speech discrimination task, the change/no-change procedure, in which a varying number of speech stimuli are presented during a trial. Standard stimuli are followed by comparison stimuli that are identical to or acoustically different from the standard. Fourteen adults and 30 4- and 5-year-old children were tested with three speech contrast pairs at a variety of signal-to-noise ratios using various numbers of standard and comparison stimulus presentations. Adult speech discrimination performance followed the predictions of the multiple looks hypothesis [N. F. Viemeister and G. H. Wakefield, J. Acoust. Soc. Am. **90**, 858–865 (1991)]; there was an increase in d by a factor of 1.4 for a doubling in the number of standard and comparison stimulus presentations near d values of 1.0. For children, increasing the number of standard stimuli improved discrimination performance, whereas increasing the number of comparisons did not. The multiple looks hypothesis did not explain the children's data. They are explained more parsimoniously by the developmental weighting shift [Nittrouer *et al.*, J. Acoust. Soc. Am. **101**, 2253–2266 (1993)], which proposes that children attend to different aspects of speech stimuli from adults. [Work supported by NIDCD and ASHF.]

5pSC2. Acoustic correlates of the question–statement contrast in children. Rupal Patel, Mariam Syeda (Dept. of Speech Lang. Pathol. and Audiol., Northeastern Univ., 360 Huntington Ave., Boston, MA 02115, r.patel@neu.edu), and Maria Grigos (New York Univ., New York, NY 10003)

The acoustics of prosodic control in children was studied in 4-, 7-, and 11-year olds using the question–statement contrast. Each child produced the utterances "Show Bob a bot" (voiced consonants) and "Show Pop a pot" (voiceless consonants) ten times each as a question and ten times each as a statement. A total of 40 utterances were analyzed per child. The following acoustic measures were obtained for each word within each utterance: average fundamental frequency ($f0$), peak $f0$, slope of $f0$, average intensity, peak intensity and duration. Preliminary results indicate no significant difference between questions and statements for 4-year olds in both the voiced and voiceless consonant conditions. In contrast, 7- and 11-year olds differentiated questions from statements by increasing average, peak, and slope of f0 as well as increasing the duration of the final syllable. Changes in syllable duration between questions and statements were more pronounced for the utterance with voiced consonants. These findings suggest that the acoustics of prosodic contrasts begins to differentiate somewhere between ages 4 and 7 and is influenced by developmental changes in physiological control and flexibility which may also affect segmental features.

5pSC3. Preference patterns in infant vowel perception. Monika T. Molnar and Linda Polka (1266 Pine Ave. W., Montreal, QC H3G 1A8, Canada, linda.polka@mcgill.ca)

Infants show directional asymmetries in vowel discrimination tasks that reveal an underlying perceptual bias favoring more peripheral vowels. Polka and Bohn (2003) propose that this bias is language independent and plays an important role in the development of vowel perception. In the present study we measured infant listening preferences for vowels to assess whether a perceptual bias favoring peripheral vowels can be measured more directly. Monolingual (French and English) and bilingual infants completed a listening preference task using multiple natural tokens of German /dut/ and /dyt/ produced by a male talker. In previous work, discrimination of this vowel pair by German-learning and by English-learning infants revealed a robust directional asymmetry in which /u/ acts as a perceptual anchor; specifically, infants had difficulty detecting a change from /u/ to /y/, whereas a change from /y/ to /u/ was readily detected. Preliminary results from preference tests with these stimuli show that most infants between 3 and 5 months of age also listen longer to /u/ than to /y/. Preference data obtained from older infants and with other vowel pairs will also be reported to further test the claim that peripheral vowels have a privileged perceptual status in infant perception.

5pSC4. Acoustical study of the development of stop consonants in children. Annika K. Imbrie (MIT, 77 Massachusetts Ave., Rm. 36-545, Cambridge, MA 02139, imbrie@mit.edu)

This study focuses on the acoustic patterns of stop consonants and adjacent vowels as they develop in young children (ages 2.6–3.3) over a 6-month period. The acoustic properties that are being measured for stop consonants include spectra of bursts, frication noise and aspiration noise, and formant movements. Additionally, acoustic landmarks are labeled for measurements of durations of events determined by these landmarks. These acoustic measurements are being interpreted in terms of the supraglottal, laryngeal, and respiratory actions that give rise to them. Preliminary data show that some details of the child's gestures are still far from achieving the adult pattern. The burst of frication noise at the release tends to be shorter than adult values, and often consists of multiple bursts, possibly due to greater compliance of the active articulator. From the burst spectrum, the place of articulation appears to be normal. Finally, coordination of closure of the glottis and release of the primary articulator is still quite variable, as is apparent from a large standard deviation in VOT. Analysis of longitudinal data on young children will result in better models of the development of motor speech production. [Work supported by NIH Grants DC00038 and DC00075.]

5pSC5. The time course of laryngeal coarticulation in children: First results. Laura L. Koenig (Long Island Univ., Brooklyn, NY and Haskins Labs., 270 Crown St., New Haven, CT 06511)

Previous work has suggested that the degree or extent of coarticulation is more extreme in young children than adults. Research in this area has focused primarily on supralaryngeal aspects of coarticulation, using spectral measures such as formant frequencies and fricative centroids. At the same time, studies of adults have found that laryngeal adjustments for voiceless consonants extend well into neighboring vowels, yielding higher values of open quotient and DC flow, and more symmetical pulse shapes, in vowels flanking voiceless as compared to voiced consonants. The current work investigates laryngeal coarticulation in normally developing, English-speaking 4- and 5-year olds. Inverse filtering of the oral airflow is used to approximate the glottal source signal in utterances containing /VpV, VhV, VbV/ sequences. The /VbV/ utterance, which does not require vocal-fold abduction, serves as a control condition to the voiceless consonants. Voice source (open quotient, speed quotient) and aerodynamic (AC and DC flow) quantities are measured over time, and compared between the children and adult females. Along with adding to our understanding of developmental changes in coarticulation, these data will contribute to the literature on differences in voice source properties between children and adults. [Work supported by NIH.]

5pSC6. Fundamental frequency and intensity variability in young children's speech produced at comfortable effort level. William S. Brown, Jr. and Rahul Shrivastav (Dept. of Commun. Sci. and Disord., Univ. of Florida, Dauer Hall, P.O. Box 117420, Gainesville, FL 32611)

Experiments to study speech production and clinical evaluation of speech often require that participants phonate or produce speech at a comfortable level of effort. It is necessary to know the degree of variability in fundamental frequency ($F0$) and intensity (SPL) across multiple test-sessions when speech is elicited in such a manner. If the intersession variability in these measures is too large, the experimental task may need to be appropriately modified. Although the variability in $F0$ and SPL produced at comfortable effort level across multiple sessions has been reported for normal adult speakers, no such data is available for young children. A vowel, a sentence, and four words were produced at a comfortable level by 15 males and 14 females between the ages of 3 and 4 years, on three separate days, one week apart. The $F0$ and SPL for these tasks were compared across the three test sessions. Results show that there were no significant differences in $F0$ and SPL and suggest that normal-speaking children between the ages of 3 and 4 years show the same degree of variability in their $F0$ and SPL as adults when asked to set their own level of comfort.

5pSC7. A developmental acoustic characterization of English diphthongs. Sungbok Lee (Depts. of Elec. Eng. and Linguist., USC, 3601 Watt Way, GFS 301, Los Angeles, CA 90089-1693, sungbokl@usc.edu), Shrikanth Narayanan, and Dani Byrd (USC, Los Angeles, CA 90089)

Acoustic properties of diphthongs in American English in a /bV/ context were investigated as a function of the age of child speakers (5–18 years, 450 subjects). Duration and the average amount of spectral change per frame were computed as a measure of the contrast between initial and final steady states, and the center position of the spectral transition (normalized for vowel length) was estimated as a measure of timing between the two steady-state portions. Results indicate that duration and temporal variability are significantly larger for children below age 10. The spectral change per frame was larger for younger children, suggesting more spectral contrast between diphthong steady states. Interestingly, the normalized center position of the glide occurs at a position relatively closer to the final portion of diphthongs for /ay/ for younger children; an opposite tendency was observed in the case of /ow/, implying that younger children exhibit delayed or premature spectral transitions. Together these observations suggest the possibility that 5–8-year olds may not produce diphthongs as a unit but, rather, as a concatenation of two vowel targets in which one component is more prolonged than the other and abrupt transitions occur. [Work supported by SBC/TRI, NSF, and NIH.]

5pSC8. Spectral moments versus Bark cepstrum classification of children's voiceless stops. James Polikoff, Jenna Hammond, Jane McNicholas, and H. Timothy Bunnell (Speech Res. Lab, duPont Hospital for Children, 1600 Rockland Rd., Wilmington, DE 19807)

Spectral moments have been shown to be effective in deriving acoustic features for classifying voiceless stop release bursts [K. Forrest, G. Weismer, P. Milenkovic, and R. N. Dougall, J. Acoust. Soc. Am. **84**, 115–123 (1988)]. In this study, we compared the classification of stops /p/, /t/, and /k/ based on spectral moments with classification based on an equal number of Bark cepstrum coefficients. The speech tokens were 446 instances each of utterance-initial /p/, /t/, and /k/ sampled from utterances produced by 208 children 6 to 8 years old. Linear discriminant analysis (LDA) was used to classify the three stops based on four analysis frames from the initial 40 ms of each token. The best classification based on spectral moments used all four spectral moment features and all four time intervals and yielded 75.6% correct classification. The best classification based on Bark cepstrum yielded 83.4% correct also using four coefficients and four time frames. Differences between these results and previous classification results using spectral moments will be discussed. Implications for future research on the acoustic characteristics of children's speech will be considered.

5pSC9. Acoustic characterization of developmental speech disorders. H. Timothy Bunnell, James Polikoff, Jane McNicholas, Rhonda Walter (Speech Res. Lab, duPont Hospital for Children, 1600 Rockland Rd., Wilmington, DE 19807, bunnell@asel.udel.edu), and Matthew Winn (Univ. of Delaware)

A novel approach to classifying children with developmental speech delays (DSD) involving /r/ was developed. The approach first derives an acoustic classification of /r/ tokens based on their forced Viterbi alignment to a five-state hidden Markov model (HMM) of normally articulated /r/. Children with DSD are then classified in terms of the proportion of their /r/ productions that fall into each broad acoustic class. This approach was evaluated using 953 examples of /r/ as produced by 18 DSD children and an approximately equal number of /r/ tokens produced by a much larger number of normally articulating children. The acoustic classification identified three broad categories of /r/ that differed substantially in how they aligned to the normal speech /r/ HMM. Additionally, these categories tended to partition tokens uttered by DSD children from those uttered by normally articulating children. Similarities among the DSD children and average normal child measured in terms of the proportion of their /r/

productions that fell into each of the three broad acoustic categories were used to perform a hierarchical clustering. This clustering revealed groupings of DSD children who tended to approach /r/ production in one of several acoustically distinct manners.

5pSC10. Learning to talk. Piers Messum (Dept. of Phonet., Univ. College, London, UK)

Is imitation a necessary part of learning to talk? The faithful replication by children of such arbitrary phenomena of English as tense and lax vowel properties, ''rhythm,'' and context-dependent VOT's seems to insist that it is. But a nonimitative account of this is also possible. It relies on two principal mechanisms. First, basic speech sounds are learned by emulation: attempting to reproduce the results achieved by other speakers but without copying their actions to do so. The effectiveness of the output provides sufficient feedback to inform the child of the adequacy of its performance and to guide refinement. Second, phonetic phenomena such as those above appear through aerodynamic accommodation. Key elements of this are (a) that speech breathing is a complex motor skill which dominates other articulatory processes during acquisition and starts pulsatile before becoming smooth, and (b) that a child-scale production system imposes constraints on talking which do not operate in the adult speaker. Much of ''the terrible complexity of phonetic patterns'' [J. Pierrehumbert, Lang. Speech **46**, 115–154 (2003)] then becomes epiphenomenal: appearing not as a result of young learners copying phonetic detail that is not linguistically significant, but of them reconciling conflicting production demands while just talking.

5pSC11. Exploring production-perception relationships for 4-year-old children: A study of compensation strategies to a lip-tube perturbation. Lucie Menard (Departement de linguistique et de didactique des langues, UQAM, CP 8888, succ. Ctr.-Ville, Montreal, QC H3C 3P8, Canada, menard.lucie@uqam.ca), Pascal Perrier, and Christophe Savariaux (CNRS-INPG-Universite Stendhal, Grenoble, France)

The relationships between production and perception for 4-year-old children were examined through a study of compensation strategies to a lip-tube perturbation. Acoustic and perceptual analyses of the rounded vowel [u] produced by 12 4-year-old French speakers were conducted under two conditions: in normal condition and with a 15-mm-diam tube inserted between the lips. Recordings of isolated vowels were made in normal condition before any perturbation ($N1$), immediately upon insertion of the tube ($P1$), for each of the next 20 trials in this perturbed condition ($P2$), and in normal condition after the perturbed trials ($N2$). Results of the acoustic analyses reveal speaker-dependent alteration of $F1$, $F2$, and/or $F0$ in the perturbed conditions and after the removal of the tube. For some subjects, the tube introduced very little changes; for some others, an increase of $F2$ was observed in $P1$, which was generally at least partly compensated during the $P2$ repetitions. Perceptual data are used to determine optimal combinations of $F0$, $F1$, and $F2$ (in bark) related to these patterns. The data are compared to a previous study conducted with adults [Savariaux *et al.*, J. Acoust. Soc. Am. **106**, 381–393 (1999)].

5pSC12. Identification and discrimination of Spanish front vowels. Isabel Castellanos and Luis E. Lopez-Bascuas (Universidad Complutense de Madrid, Facultad de Psicologa, Campus de Somosaguas, 28223 Madrid, Spain, lelopezb@psi.ucm.es)

The idea that vowels are perceived less categorically than consonants is widely accepted. Ades [Psychol. Rev. **84**, 524–530 (1977)] tried to explain this fact on the basis of the Durlach and Braida [J. Acoust. Soc. Am. **46**, 372–383 (1969)] theory of intensity resolution. Since vowels seem to cover a broader perceptual range, context-coding noise for vowels should be greater than for consonants leading to a less categorical performance on the vocalic segments. However, relatively recent work by Macmillan *et al.* [J. Acoust. Soc. Am. **84**, 1262–1280 (1988)] has cast doubt on the assumption of different perceptual ranges for vowels and consonants even though context variance is acknowledged to be greater for the former. A possibility is that context variance increases as number of long-term phonemic categories also increases. To test this hypothesis we focused on Spanish as the target language. Spanish has less vowel categories than English and the implication is that Spanish vowels will be more categorically perceived. Identification and discrimination experiments were conducted on a synthetic /i/-/e/ continuum and the obtained functions were studied to assess whether Spanish vowels are more categorically perceived than English vowels. The results are discussed in the context of different theories of speech perception.

5pSC13. Classification of stop consonant place of articulation. Atiwong Suchato (Speech Commun. Group, MIT, Cambridge, MA 02139, atiwong@mit.edu)

In this study we develop an experimental procedure for examining the relative importance of knowledge-based cues for identifying place of articulation for stop consonants. A set of acoustic attributes is selected for place classification of stops: amplitude and energy of burst, formant movement of adjacent vowels, spectrum of noise after the release, and some temporal cues. The ability of each attribute to separate the three places is evaluated by the classification error based on the distributions of its values for the three places, and another quantifier based on F ratio. These two quantifiers generally agree and show how well each individual attribute separates the three places. Linear discriminant function analysis is used to address the relative importance of these attributes when combinations are used. Their discriminating abilities and the ranking of their relative importance to the classification in different vowel and voicing contexts are reported. The overall findings are that attributes relating to the burst spectrum in relation to the vowel contribute most effectively, while formant transition is somewhat less effective. The approach used in this study can be applied to different classes of sound, as well as stops in different noise environments. [Work supported by NIH Grant Number DC 02978.]

5pSC14. Role of experience in eliciting perceptual overshoot with sine waves. Radhika Aravamudhan and John W. Hawks (Kent State Univ., Kent, OH 44242, raravam1@kent.edu)

Previous studies in vowel perception have demonstrated perceptual overshoot (PO) with vowels in CVC contexts. The first experiment investigated the elicitation of PO with speech and nonspeech signals. Boundary estimates were obtained for four continua: (1) steady-state synthetic vowels; (2) synthetic vowel transitions in /wV/ context; (3) sine wave analogs mimicking the steady-state vowels; and (4) sine wave analogs mimicking the transition continuum. Results replicated the findings of previous studies by demonstrating PO with synthetic vowels; however, PO was not elicited with the sine wave analogs. This led to a second experiment, where the role of training on the elicitation of PO with sine wave analogs was investigated. Subjects from the first experiment were divided into two groups. One group was trained to categorize the steady-state sine wave analogs based on their synthetic vowel boundary. Training occurred over 3–6 1-h sessions until subjects achieved 90%-correct identification. The control group received no training. Post-training boundary estimates were obtained for the four original continua from both groups. The results indicated that PO was now elicited with the sine wave analogs for the trained group, but not for the control group. Further implications of these findings will be discussed.

5pSC15. The perceptual magnet effect reflects phonetic context.
Sarah Hawkins (Dept. of Linguist., Univ. of Cambridge, Sidgwick Ave., Cambridge CB3 9DA, UK, sh110@cam.ac.uk) and Sarah Barrett Jones (City Univ., London EC1V 0HB, UK)

Two experiments demonstrate that the perceptual magnet effect is context sensitive. In experiment 1a, 24 participants rated goodness of synthetic /u/ in isolation (*ooh*) and in two consonantal contexts, /lu/, /ju/ (*Lou, you*), with nine versions per word, varying in $F2$ frequency. Their most (prototypical) and least (nonprototypical) preferred choices reflected expected differences between words, and individual differences within words. Experiment 1b demonstrated standard perceptual magnet effects for prototype and nonprototype /u/ in the three words. Unlike previous work, each participant discriminated his/her own prototype and nonprototype from experiment 1a, rather than the group mean. Experiment 2a replicated 1a with 40 new participants. Experiment 2b compared discrimination around participants' prototypical $F2$ frequency for /u/ in one word (original) with discrimination around that same frequency in another word (transferred). Different original/transferred sets were heard by four groups (ten participants each): /u/ and /lu/; /lu/ and /u/; /ju/ and /u/; /ju/ and /lu/. Discrimination (d') near the prototype was poorer for original than transferred contexts [for the four comparisons, $t(9)$ ranged between 2.43–3.49, $p<0.025$–0.005]. Thus, the perceptual magnet effect is syllable specific: the vowel prototype for one word need not generalize to another. Implications for perceptual coherence and phonological representation are discussed.

5pSC16. Phonotactic constraints, frequency, and legality in English onset-cluster perception.
Elliott Moreton (Dept. of Linguist., Univ. of North Carolina, Dey Hall, Rm. 318, Chapel Hill, NC 27599-3155, moreton@unc.edu)

Phonological context can affect phoneme identification, favoring one response over another [Massaro and Cohen, Percept. Psychophys. **34**, 338–348 (1983)]. It is unclear what is disfavored: infrequency, legality, or phonotactic constraint violation. This study compared bias on an [l]–[w] continuum in two situations. In one, the three factors were deliberately confounded; in the other, frequency and legality were controlled. Stimuli were synthetic syllables ambiguous between [blae bwae dlae dwae] ("bd array") or [mlae mwae nlae nwae] ("mn array"). In the bd array, [bl dw] are legal, frequent English onsets, while [dl bw] have zero frequency, are illegal, and violate a constraint against same-place onsets; hence, "d" responses should facilitate "w" responses. The whole mn array is unattested and illegal, but [nl mw] violate the constraint. Simultaneous stop-sonorant judgments were obtained from English listeners for each 6-by-6 array. A mixed-effects logistic-regression model (including a term to model out compensation for coarticulation) was used to measure dependency between stop and sonorant responses. The bd array drew the expected bias against "dl." Bias against "nl" in the mn array was considerably smaller, but still present. These results suggest that perceptual bias is jointly determined by constraint violation, and by frequency and/or legality.

5pSC17. Effect of retroflex sounds on the recognition of Hindi stops.
Amita Dev, S. S. Agrawal, and D. Roy Choudhary (Ambedkar Inst. of Technol., Shakarpur, Madhuban, Delhi 92, India, amita_dev@hotmail.com)

As development of the speech recognition system entirely depends upon the spoken language used for its development and the very fact that speech technology is highly language dependent and reverse engineering is not possible, there is an utmost need to develop such systems for Indian languages. In this paper we present the implementation of a time-delay neural network system (TDNN) in a modular fashion by exploiting the hidden structure of previously phonetic subcategory network for the recognition of Hindi consonants. For the present study we have selected all the Hindi phonemes for the recognition. A vocabulary of 207 Hindi words was designed for the task-specific environment and used as a database. For the recognition of phonemes a three-layered network was constructed and the network was trained using the backpropagation learning algorithm. Experiments were conducted to categorize the Hindi voiced and unvoiced stops, semivowels, vowels, nasals, and fricatives. A close observation of the confusion matrix of Hindi stops revealed maximum confusion of retroflex stops with their nonretroflex counterparts.

5pSC18. Perception of correlations between acoustic cues in category tuning and speaker adaptation.
Lori Holt and Travis Wade (Dept of Psych., Carnegie Mellon Univ., 5000 Forbes Ave., Pittsburgh, PA 15213, lholt@andrew.cmu.edu)

In English and many other languages, fundamental frequency $(f0)$ varies with voicing such that voiced consonants are produced with lower $f0$'s than their voiceless counterparts. This regularity robustly influences perception, such that sounds synthesized or spoken with a low $f0$ are more often perceived as voiced than are sounds with a higher $f0$. This series of studies exploited these observations to investigate category tuning as a function of incidental exposure to correlations among speech cues and adaptation to speaker idiosyncrasies or accent. Manipulation of $f0$ across sets of natural speech utterances produced stimulus sets varying in their inherent $f0$/voicing relationship. Listeners were exposed to these different $f0$/voicing patterns via spoken word and nonword items in a lexical decision task, and their resulting categorization of ambiguous consonants varying in $f0$ and voice onset time (VOT) was measured. The results suggest listeners adapt quickly to speaker-specific cues but also remain influenced by more global, naturally occurring covariance patterns of $f0$ and voicing in English. This pattern contrasts somewhat with studies where idiosyncrasy is represented instead by manipulation of primary, first-order cues to speech sounds, in which listeners are seen to adapt more straightforwardly to the cues they are presented.

5pSC19. On the categorical nature of Korean /pk/ place assimilation.
Minjung Son (Linguist. Dept., Yale Univ. and Haskins Labs., 370 Temple St., New Haven, CT 06520-8366) and Alexei Kochetov (Simon Fraser Univ., Burnaby, BC, Canada)

Korean exhibits regressive place assimilation in /pk/ clusters, which has been described as gradient and rate dependent. However, this assumption has empirically only been tested on the basis of air pressure data [Jun, 1996] which does not provide a direct record of articulator movement. The present study examines articulator movement using EMMA. For three Seoul-dialect speakers, stimuli containing /pk/ clusters were elicited word-medially (for words and nonwords) and in a phrase-boundary condition; two rates were employed. Results show that the labial can indeed reduce word medially, rendering [kk]. However, contrary to previous claims, the data demonstrate that reduction in /pk/ is always categorical, although it is optional or stochastic in its occurrence. Substantial interspeaker variation is observed, with the frequency of reduction being higher at fast rate and ranging overall from 6 at both rates and is never gradient. The lack of reduction in nonsense words and in the phrase boundary condition shows that the process is sensitive to lexical properties. The observed tendency for more gestural overlap word medially compared to the phrase-boundary condition supports the hypothesis that gestural overlap plays a role in the origins of place assimilation. [Work supported by NIH.]

5pSC20. Information conveyed by vowels about other vowels.
Hector R. Javkin, Elaine Drom, Carol Christie, and Gaston R. Cangiano (Dept. of Linguist. and Lang. Development, San Jose State Univ.)

Rapid adaptation to different speakers has become an important issue in speech recognition (ASR) but has been known in human listeners for a long time. Ladefoged and Broadbent [J. Acoust. Soc. Am. **29**, 98–104 (1957)] demonstrated that human perception of synthesized vowels occurring in monosyllabic words (bit, bet, bat, but) can be changed by changing the formants of an introductory phrase (Please say what this word is).

These stimuli meant that the introductory phrase ranged over the same portion of the vowel space (front vowels, or relatively high $F2$), thus facilitating listeners adaptation. To further test the limits of human adaptation, we replicated the experiment keeping the same words, substituting introductory phrases consisting of back (low $F2$) vowels, and maintaining a similar level of low-quality synthesized speech. The effects are difficult to replicate with natural or high-quality synthetic speech. However, we will suggest that low quality speech is analogous to the low-dimensionality representation of speech of many ASR front ends, which discard, for example, information as to the higher formants. Therefore, these findings are relevant to the use of adaptation in improving ASR. [Work supported by a Faculty Small Grant from San Jose State University.]

5pSC21. On the relation of apparent naturalness to phonetic perceptual identification of vowels. Stephanie Wissig, Daria F. Ferro, Kate Liberman, Jill Thompson, and Robert E. Remez (Dept. of Psych., Barnard College, 3009 Broadway, New York, NY 10027)

A set of synthetic test syllables was created varying the attributes of naturalness and phonetic vowel height and advancement. These acoustic items were used in tests calibrating the relation of naturalness to phonetic perceptual resolution. Two acoustic methods were used to create naturalness variants: (1) variation in the excitation of the synthetic voicing source and (2) variation in the bandwidth of the formant centers. A naturalness tournament was composed of items drawn from the test series, and the sensitivity of perceivers to the vowel contrast was estimated with the cumulative d' across the series in identification tests. These outcomes reveal different effects of each acoustic manipulation: an independence of the outcome of the naturalness tournament and the measures of phonetic sensitivity in one case and a close relation between naturalness and phonetic resolution in the other. Together, the findings show that intelligibility and naturalness can be either orthogonal or contingent aspects of speech perception. These measures offer a tool to understand rule-based and exemplar-based components of phonetic perception. [Research supported by NIH (DC00308).]

5pSC22. Spectral frequency modulation in vowel identification. Chang Liu and David A. Eddins (Psychoacoustic Lab., Dept. of Commun. Disord. and Sci., Univ. at Buffalo, Buffalo, NY 14214)

Psychophysical and physiological studies have demonstrated selectivity for spectral envelope frequency (also termed spatial frequency) in the auditory system, suggesting that auditory perception of complex sounds might be based on spectral envelope channels. The present study investigated relative contribution of different spatial frequencies to vowel identification. Twelve naturally-spoken American-English vowels were presented at 70 dB SPL. In different conditions, vowel stimuli were subjected to various degrees of low-pass and high-pass filtering in the spatial frequency domain, in effect, altering their spectra. Identification performance for the vowels with and without spatial frequency filtering was estimated for normal-hearing listeners. Results indicated that vowel identification performance was progressively degraded as spatial-frequency components were removed. Results will be interpreted in terms of spatial frequency regions most important to specific vowel categories. The specificity and universality of spatial frequency modulations in vowel identification across different vowel categories will be discussed.

5pSC23. Gradient sensitivity to acoustic detail and temporal integration of phonetic cues. Bob McMurray, Meghan A. Clayards, Richard N. Aslin, and Michael K. Tanenhaus (Univ. of Rochester, Rochester, NY 14627)

Speech contains systematic covariation at the subphonemic level that could be used to integrate information over time (McMurray *et al.*, 2003; Gow, 2001). Previous research has established sensitivity to this variation: activation for lexical competitors is sensitive to within-category variation in voice-onset-time (McMurray *et al.*, 2002). This study extends this investigation to other subphonemic speech cues by examining formant transitions (r/l and d/g), formant slope (b/w) and VOT (b/p) in an eye-tracking paradigm similar to McMurray *et al.* (2002). Vowel length was also varied to examine temporal organization (e.g., VOT *precedes* the vowel). Subjects heard a token from each continua and selected the target from a screen containing pictures of the target, competitor and unrelated items. Fixations to the competitor increased with distance from the boundary along each of the speech continua. Unlike prior work, there was also an effect on fixations to the *target*. There was no effect of vowel length on the d/g or r/l continua, but rate dependent continua (b/w and b/p) showed length effects. Importantly, the temporal order of cues was reflected in the pattern of looks to competitors, providing an important window into the processes by which acoustic detail is temporally integrated.

5pSC24. A novel acoustic way of demarcating diphthong elements in Thai. Rungpat Roengpitya (Dept. of Linguist., Faculty of Arts, Chulalongkorn Univ., Bangkok 10330, Thailand, rungpatr@hotmail.com)

Thai has three phonemic (falling) diphthongs /ia/, /iʻa/, and /ua/ and a set of the so-called rising diphthongs in Thai /ai/ /aj/ and /au/ /aw/. In previous (nonacoustic) literature, falling diphthongs are short when followed by a glottal stop. However, recent acoustic studies revealed that falling diphthongs in Thai are short in closed syllables, but long in open syllables. In the same literature, each diphthong was measured for the durations of its first vocalic element, the transition, and the second vocalic element, without demarcating the offset of the first vocalic element and the onset of the second vocalic element. This paper has the major aim of applying a new technique for making a plausible way to mark where the first vocalic element ends and the second vocalic element starts, not only for the set of falling diphthongs but also for the so-called rising diphthongs. Moreover, this new technique would help to find out whether the rising diphthongs in Thai have the combinations of two-vowel qualities or a vowel and a glide. It is hoped that this new method will be useful for the future acoustic studies in other languages of the world. [Work supported by new-lecturer grants, Chulalongkorn University.]

5pSC25. Obtaining a palatal trace for ultrasound images. Melissa A. Epstein, Maureen Stone, Marianne Pouplier (Vocal Tract Visualization Lab, Biomed. Sci., Univ. of Maryland Dental School, Rm. 5A12, 666 W. Baltimore St., Baltimore, MD 21201, mae001@dental.umaryland.edu), and Vijay Parthasarathy (Johns Hopkins Univ., Baltimore, MD 21218)

This paper presents methods for collection and display of the palate with ultrasound, for use as a reference for tongue movements. Ultrasound does not usually capture structures other than the tongue, because the air above the tongue in the vocal tract reflects the ultrasound beam back to the transducer. However, when the tongue touches the palate, the ultrasound beam is transmitted through the soft tissue until it reaches and is reflected by the palatine bone. The tongue touches the palate during swallowing and some speech sounds. The palate contour can be traced from these images. The paper presents a corpus of speech and swallowing tasks that can be used to create a full palatal trace. The corpus is tested on a subject for whom it is easy to collect palatal images and a subject for whom it is difficult to collect palatal images. The availability of a palate will enhance our ability for data quantification from ultrasound images. In combination with tongue contours, the palate contour allows the computation of lin-

guistically important measures, such as the constriction degree, area functions, and $L2$ norms. [Work supported by NIH RO1-DC01758 and T32-DE07309.]

5pSC26. Effects of human fatigue on speech signals. Catherine Stamoulis (Cambridge Sci. Solutions, 140 Marlborough St., Boston, MA 02116)

Cognitive performance may be significantly affected by fatigue. In the case of critical personnel, such as pilots, monitoring human fatigue is essential to ensure safety and success of a given operation. One of the modalities that may be used for this purpose is speech, which is sensitive to respiratory changes and increased muscle tension of vocal cords, induced by fatigue. Age, gender, vocal tract length, physical and emotional state may significantly alter speech intensity, duration, rhythm, and spectral characteristics. In addition to changes in speech rhythm, fatigue may also affect the quality of speech, such as articulation. In a noisy environment, detecting fatigue-related changes in speech signals, particularly subtle changes at the onset of fatigue, may be difficult. Therefore, in a performance-monitoring system, speech parameters which are significantly affected by fatigue need to be identified and extracted from input signals. For this purpose, a series of experiments was performed under slowly varying cognitive load conditions and at different times of the day. The results of the data analysis are presented here.

5pSC27. HOCUS: The Haskins optically-corrected ultrasound system for measuring speech articulation. D. H. Whalen, Khalil Iskarous, Mark K. Tiede, and David J. Ostry (Haskins Labs., 270 Crown St., New Haven, CT 06511)

The tongue is the most important supralaryngeal articulator for speech, yet, because it is typically out of view, its movements have been difficult to quantify. Here is described a new combination of techniques involving ultrasound in conjunction with an optoelectric motion measurement system (Optotrak). Combining these, the movements of the tongue are imaged and simultaneously corrected for motion of the head and of the ultrasound transceiver. Optotrak's infrared-emitting diodes are placed on the transceiver and the speakers head in order to localize the ultrasound image of the tongue relative to the hard palate. The palate can be imaged with ultrasound by having the ultrasound signal penetrate a water bolus held against the palate by the tongue. This trace is coregistered with the head and potentially with the same talker's sagittal MR image, to provide additional information on the unimaged remainder of the tract. The tongue surface, from the larynx to near the tip, can then be localized in relationship to the hard palate. The result is a fairly complete view of the tongue within the vocal tract at sampling rates appropriate for running speech. A comparison with other imaging vocal tract systems will be presented. [Work supported by NIH Grant DC-02717.]

5pSC28. Movie MRI at five frames a second for evaluation of speech and swallowing. Masanobu Kumada (Dept. of Otorhinolaryngol., Hospital, Natl. Rehabilitation Ctr. for the Disabled (NRCD), 4-1 Namiki, Tokorozawa, Saitama 359-8555, Japan), Koichi Mori, Yasoichi Nakajima (NRCD, 4-1 Namiki, Tokorozawa, Saitama 359-8555, Japan), and Seiji Nozaki (Toshiba Medical Systems Corp.)

Magnetic resonance imaging (MRI) is a noninvasive imaging method that is widely used in the medical field. One of the limitations of MRI is its low time-resolution; images of MRI are usually obtained as still images. Here we introduced a newly developed method of "movie" MRI with high time-resolution at five images a second. Its good application would include study and evaluation of speech and swallowing. Instrument: MRT-2001 XG with Software ver.5.5. (Toshiba). Coil: head QD coil for the tongue; CTL ARRY coil (3ch) for the neck, Imaging condition: sagital FFE2D; TR$=2.8$ ms; TE$=1.2$ ms; FA$=10$ deg; Matrix$=64\times128$; ST $=10$ mm; NAQ$=1$ AV; FOV$=23\times35$; ROAFI; sequential acquisition. Maximum imaging length$=51$ s. A healthy Japanese male (Tokyo dialect speaker, 39 years). Task: Repetitive utterance of /tenten . . . / and intentional swallowing of saliva. Results: In the task of /tenten . . . ./, we could detect, in temporal order, velopharyngeal (VP) closure, opening of the tongue-palate (TP) closure, VP opening, TP closure, VP closure, and so on. In the swallowing task of saliva, we could detect movement of the tongue for conveying saliva to the pharynx, VP closure, backward movement of the tongue root, elevation and descent of the larynx in this temporal order. Our "Movie MRI" seemed promising for the noninvasive evaluation of speech and deglutition. Appropriate materials for swallowing evaluation will be presented.

5pSC29. Anterior tongue and jaw movement in sVd words. Richard S. McGowan (CReSS LLC, 1 Seaborn Pl., Lexington, MA 02420, cressllc@earthlink.net)

The relations between jaw and tongue movements were examined using words of the form sVd for eight different speakers from the X-Ray Microbeam Speech Production Database. Measurements were examined at the maximum speeds during the release and during the closure of the tongue blade and tongue body. For nonhigh vowels the tongue blade traveled in the same direction as the jaw during release, and, to a lesser extent, the same was true during closure. Further, the magnitude of the projection of the tongue blade velocity onto the direction of the jaw movement was often large compared with the speed of the jaw. There was less consistency in the relation between tongue body and jaw movement. These results indicate that the jaw and tongue movements are not rigidly coupled. Rather the jaw, which can provide a hard boundary for the tongue, is getting out of the tongue's way during release and following the tongue on closure for subsequent bracing. [Work supported by Grant NIDCD-01247 to CReSS LLC.]

5pSC30. Tongue–jaw kinematic correlates of /s/ spectra. James S. Dembowski and Richard K. Crumb (Commun. Disord., State Univ. of New York at New Paltz, HUM 14a, 75 S. Manheim Blvd., New Paltz, NY 12561-2499, dembowsj@newpaltz.edu)

Frequencies of spectral peaks for fricatives are determined by the size of the resonating cavity anterior to the place of articulatory constriction in the upper vocal tract. For /s/, this cavity size may be altered through anterior–posterior (a–p) movements of the tongue blade forming the constriction, changes of jaw height, and degree of lip protrusion. With respect to intensity, modeling studies suggest that intensity of fricative spectral peaks may be related to degree of articulatory constriction. These spectral–kinematic relationships have been little studied in natural speech. This study used data from the University of Wisconsin X-Ray Microbeam Database to examine the relationship between spectral peaks and movements of the tongue and jaw in the /s/ productions of one normal speaker. Results showed no relationship between a–p tongue position and frequency of spectral peaks. However, a significant inverse correlation related peak between frequency and jaw opening. Thus, for this speaker jaw height appeared a more important determinant of spectral variability for /s/ than tongue position. Additional results showed a significant relationship between peak intensity and distance of the tongue blade from the palate. These natural speech data will be discussed with respect to models and theories of fricative production.

5pSC31. Correlation between angle of incidence and sliding patterns of the tongue along the palate in Korean velar stops. Jana Brunner (Zentrum fuer Allgemeine Sprachwissenschaft [ZAS] Berlin, Germany, brunner@zas.gwz-berlin.de), Susanne Fuchs (ZAS, Berlin, Germany), Pascal Perrier (Institut de la Commun. Parle, & Univ. Stendhal, Grenoble, France), and Hyeon_Zoo Kim (Dankook Univ., Seoul, Korea)

In former studies, it has been hypothesized that the articulatory production of oral stops could result from the interaction between the tongue moving towards a virtual target located above the palate, and the palate. Velar stops, where the tongue slides along the palate during the occlusion phase, offer a nice experimental framework for further experimental assessments. Indeed, in the framework of the ''virtual target'' hypothesis, the sliding movement should be seen as the continuation of the movement before the occlusion, but constrained by the palate. Hence, relations should exist between the movement characteristics before contact and during the occlusion phase. To test this hypothesis three Korean speakers were recorded via EMA producing /aCV/ sequences with C=/g/, /k'/ and /kh/, V=/a/, /i/ or /u/. The angle between tongue trajectory just before the impact and palatal contour was estimated, and the amplitude of the sliding movement was measured. Preliminary results for two speakers show that these two variables correlate: The greater the angle, the larger the sliding movement. These findings are interpreted as supporting the ''virtual target'' hypothesis. This interpretation will be verified by simulations using a 2D biomechanical tongue model [Payan and Perrier, Speech Commun. **22** (1997)].

5pSC32. Modeling the acoustics of American English /r/ using configurable articulatory synthesis (CASY).
Heike Lehnert-LeHouillier, Khalil Iskarous, and Douglas H. Whalen (Haskins Labs., 270 Crown St., New Haven, CT 06511)

The claim that articulatory variation in /r/ production exhibits systematic tradeoffs to achieve a stable acoustic signal (Guenther *et al.*, 1999) was tested using configurable articulatory synthesis (CASY) and ultrasound data. In particular, the hypothesis was tested that multiple constrictions during /r/ production are necessary to achieve a low enough $F3$. Ultrasound and Optotrak data from four speakers pronouncing /r/ in different vocalic contexts were used to determine where in the vocal tract the tongue gestures are placed. This data was then modeled using CASY parameters and was used to determine how the three gestures in /r/ (labial, palatal, and pharyngeal) contribute to the $F3$ value observed in the speech signal simultaneously recorded with the ultrasound. This was done by varying the degree and location of the lingual constrictions and the degree of the labial constriction and determining the effect on $F3$. It was determined that the three gestures in /r/ contribute in differing amounts to the overall $F3$ lowering. Furthermore, it does not seem that all three gestures are necessary for $F3$ lowering. This lends support to the hypothesis that the goal in /r/ production is the simultaneous achievement of three gestures. [Work supported by NIH Grant DC-02717.]

5pSC33. Formants and musical harmonics matching in Brazilian lied.
Beatriz Raposo de Medeiros (Universidade de So Paulo Av Prof. Luciano Gualberto, 403, CEP 01060-970, Sao Paulo, Brazil)

This paper reports a comparison of the formant patterns of speech and singing. Measurements of the first three formants were made on the stable portion of the vowels. The main finding of the study is an acoustic effect that can be described as the matching of the vowel formants to the harmonics of the sung note (A flat, 420 Hz). For example, for the vowel [a], $F1$ generally matched with the second harmonic (840 Hz) and $F2$ with the third harmonic. This finding is complementary to that of Sundberg (1977) according to which the higher the fundamental frequency of the musical note, e.g., 700 Hz, the more the mandible is lowered causing the elevation of the first formant of the sung vowel. As Sundberg himself named this phenomenon, there is a matching between the first formant and the phonation frequency, causing an increase in the sound energy. The present study establishes that the matching affects not only $F1$ but also $F2$ and $F3$. This finding will be discussed in connection with other manoeuvres (e.g., tongue movements) used by singers.

5pSC34. Vowel length in Farsi.
Shabnam Shademan (Linguist. Dept., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095)

This study tests whether Farsi vowels are contrastive with respective to length. Farsi has a six-vowel system with three lax vowels and three tense vowels. Both traditional grammarians and modern linguists believe that Farsi tense vowels are longer than lax vowels, and that there are no vowel pairs that contrast only in length. However, it has been suggested that Farsi exhibits compensatory lengthening, which is triggered by the deletion of glottal consonants in coda position in informal speech (Darzi, 1991). As a result, minimal pairs such as [tar] and [tarh] should contrast only with respect to vowel length. A corpus of 90 words of the form CVC, CVCG, CVGC, and CVCC (where V=a vowel and G=a glottal consonant) was recorded, and durations of vowels in different contexts were measured and compared. Preliminary results show that lax vowel durations fall into three groups with CVCC longer than CVCG/CVGC, and the latter longer than CVC. It remains to be seen whether CVCG/CVGC words show compensatory lengthening when the glottal consonant is deleted.

5pSC35. Gestural stability in vowels.
Thomas Purnell (Dept. of Linguist., Univ. of Wisconsin, Madison, 1168 Van Hise, 1220 Linden Dr., Madison, WI 53706, tcpurnell@wisc.edu)

In accordance with proper perception of linguistic sound units, past research has demonstrated some degree of acoustic and physiological stability. In contrast, articulatory stability has been thought to be inconsistent because articulations may vary so long as the vocal tract area function results in appropriate formant structure [Atal *et al.*, J. Acoust. Soc. Am. **63**, 1535–1555 (1978)]. However, if the area function for the constriction and its anterior region can maintain acoustic stability, articulatory stability should be observed in the relational behavior of four tongue pellets used in xray microbeam data. Previous work examined normalized pellet data in order to arrive at an average posture for each vowel [Hashi *et al.*, J. Acoust. Soc. Am. **104**, 2426–2437 (1998)]. But by assuming static (average) gestures, the research fell short of a correct postural characterization. This study of tongue pellet speed and normalized pellet displacement of front vowels spoken by ten microbeam database subjects reports that the tongue tip pellet speed maxima identify vowel edges (end of vowel onset, beginning of offset) while displacement of the three anterior pellets identify changes in formant structure (e.g., two stable regions in the Northern Cities English front low vowel).

5pSC36. Aeroacoustics production of fricative speech sounds.
Michael Krane (Ctr. for Adv. Information Processing, Rutgers Univ., Piscataway, NJ 08854) and Settles Gary (Penn State Univ., University Park, PA 16802)

The aeroacoustic production of fricative speech sounds is studied using schlieren imaging and acoustic measurements. The focus is the structure of the turbulent jets formed during fricative speech sound production, and how interaction of the jet with articulators affects the acoustic nature of the speech sound. Patterns of the jets formed during the articulation of both voiced and unvoiced fricatives (s and z) are shown using schlieren images. In particular, the interaction of the jet with articulators such as teeth and lips are clearly seen, and demonstrated by varying articulator positions. Pressure measurements were made in conjunction with the images using a microphone placed near the teeth and one in the farfield. The pressure measurements show the acoustic consequences of the various jet/articulator interactions, further clarifying which articulators are most important in determining the aeroacoustic source characteristics.

5pSC37. A new taxonomy of American English /r/ using MRI and ultrasound.
Mark K. Tiede (Haskins Labs, 270 Crown St., New Haven, CT 06511 and M.I.T.-R.L.E., Cambridge, MA, tiede@haskins.yale.edu), Suzanne E. Boyce, Christy K. Holland, and K. Ann Choe (Univ. of Cincinnati, Cincinnati, OH)

In this work we present preliminary results from a large scale production study of American English liquids. MRI and ultrasound have been used to image 20 subjects producing /r/ and /l/. Subjects were native speakers, representative of the main American dialects, and balanced between sexes. MRI data were collected volumetrically for recovery of

three-dimensional tongue shape. In addition, a short (1 s) midsagittal protocol was used in conjunction with ultrasound scanning to confirm the validity of the sustained production required for volumetric imaging. Our results show tongue shapes not inconsistent with the six main types found by Delattre and Freeman (1968), but their range and variety suggest that either additional canonical shapes or a different organizational principal (such as palatal shape) is motivated. We have also observed consistent patterns of tongue grooving associated with different midsagittal shapes, from the parasagittal data not accessible from the Delattre and Freeman cineradiography. In addition to this 3D-based taxonomy of /r/ tongue shapes we present similar data for lateral productions, and the associated acoustic structures for both. [Research supported by NIH.]

5pSC38. Some effects of prosodic structure on the production of /u/ in French. Marija Tabain (MARCS, UWS, Sydney, NSW 1797, Australia), Pascal Perrier, and Christophe Savariaux (ICP, INPG, Grenoble, France)

In this paper we present formant data and EMA (Carstens) data from two female speakers of French who produced /u/ in domain-final position at four different prosodic boundaries (in hierarchical order: utterance, intonational phrase, accentual phrase, and word boundaries). The prosodic boundaries are used in order to control the acoustic duration of the /u/, with greatest duration at the strongest boundary (utterance) and shortest duration at the weakest boundary (word). We present results on lip protrusion and tongue body targets with respect to these prosodically induced differences in duration, and compare these articulatory results with measurements of $F1$ and $F2-F0$ in Bark. We examine trade-offs between lip and tongue targets as duration is reduced, and speculate as to whether the speakers may have been aiming for an articulatory target or an acoustic target for /u/ at the stronger prosodic boundaries. This work combines our previous work on articulatory prosody of /a/ [Tabain, J. Acoust. Soc. Am. **113**, 516–531 (2003)] with our previous work on perturbation of /u/ [Savariaux *et al.*, J. Acoust. Soc. Am. **98**, 2428–2442 (1995)]. [Work supported by an Australian Research Council fellowship to the first author.]

5pSC39. Production and perception of Persian geminate stops at three speaking rates. Benjamin B. Hansen (Dept. of Linguist., Univ. of Texas, Calhoun Bldg. 501, Austin, TX 78712)

An experiment was designed to determine whether the geminate/singleton category distinction is maintained at fast speaking rates in Persian. Three speakers of Tehrani Persian read test words containing [t,t:,d,d:] in carrier sentences at three speaking rates. The categories do not overlap within a given speaking rate, but the fastest geminates do overlap the normal-rate singletons, implying that the listener must take speaking rate into account in order to perceive the category distinction. The ratio of the consonant closure to the preceding vowel (C/V) is not a useful rate-independent parameter for describing the geminate/singleton boundary in Persian since in Persian the vowel preceding a geminate is slightly longer. However, it was found that the marginal consonant closure (above a minimum closure of about 20 ms) maintains a fixed proportion of the average syllable duration, regardless of rate. This fixed proportion is distinct for geminates and singletons, and so may be used as a single rate-independent parameter for defining the category distinction. Perception tests on natural sentences showed that the distinction is perceptible at each of the three speaking rates. The perceptual response to manipulation of the closure durations indicated that, besides duration, additional cues to the distinction are present.