NOTE: All Journal articles and Letters to the Editor are peer reviewed before publication. Program abstracts, however, are not reviewed before publication, since we are prohibited by time and schedule.

MONDAY MORNING, 24 MAY 2004  ROYAL BALLROOM A, 7:40 A.M. TO 12:00 NOON

Session 1aAA

Architectural Acoustics and Noise: Effect of Room Acoustic Environment on Human Productivity and Performance

Rendell F. Torres, Cochair
Architectural Acoustics Program, Rensselaer Polytechnic Institute School of Architecture, 110 8th Street, Troy, New York 12180-3590

Lily M. Wang, Cochair
Architectural Engineering, University of Nebraska—Lincoln, 200B Peter Kiewit Institute, 1110 South 67th Street, Omaha, Nebraska 68182-0681

Chair’s Introduction—7:40

Invited Papers

7:45


It is of interest to estimate the influence of the environment in a specific work place area on the performance and well-being of people. Investigations have been carried out for the cabin environment of an airplane and for class rooms. Acoustics is only one issue of a variety of environmental factors, therefore the combined impact of temperature, humidity, air quality, lighting, vibration, etc. on human perception is the subject of psychophysical research. Methods for the objective assessment of subjective impressions have been developed for applications in acoustics for a long time, e.g., for concert hall acoustics, noise evaluation, and sound design. The methodology relies on questionnaires, measurement of acoustic parameters, ear-related signal processing and analysis, and on correlation of the physical input with subjective output. Methodology and results are presented from measurements of noise and vibration, temperature and humidity in aircraft simulators, and of reverberation, coloring, and lighting in a primary school, and of the environmental perception. [The work includes research with M. Klatte, A. Schick from the Psychology Department of Oldenburg University, and M. Meis from Hoerzentrum Oldenburg GmbH and with the European Project HEACE (for partners see www.heace.org).]

8:05

1aAA2. A review of the combined effects of thermal and noise conditions on human performance. Richard A. Moscoso, Lily M. Wang, and Amy Musser (Architectural Eng. Prog., Univ. of Nebraska–Lincoln, Peter Kiewit Inst., 1110 S. 67th St., Omaha, NE 68182-0681, moscosobullon@mail.unomaha.edu)

Human perception and annoyance due to background noise has been the subject of much research. A great deal of work has also been done to identify conditions that produce an acceptable thermal environment for building occupants. The experience of occupants in indoor environments, however, is much more complex than can be represented by thermal comfort or the acoustic environment in isolation. Occupants normally experience a mix of thermal, auditory, visual, and olfactory stimuli that combines to form an impression of the environment. This paper is specifically interested in how building occupants trade off between acoustic and thermal comfort.
Heating, ventilation, and air-conditioning systems in buildings are often adjusted by building users to arrive at a more comfortable temperature, but this change may also produce more noise. Previous studies on the interaction effects between temperature and noise on human performance are reviewed in this presentation, followed by a discussion of the authors’ current work in this area.

8:25
1aAA3. Ventilation noise and its effects on annoyance and performance. Ulf Landstrom (Natl. Inst. for Working Life North, Box 7654, SE-907 13 Umea, Sweden, ulf.landstrom@arbetslivsinstitutet.se)

In almost every room environment, ventilation acts as a more or less prominent part of the noise exposure. The contribution to the overall sound environment is a question not only of the way in which the ventilation system itself functions, but also a question of the prominence of other contemporary sound sources such as speech, equipment, machines, and external noises. Hazardous effects due to ventilation noise are most prominent in offices, hospitals, control rooms, classrooms, conference rooms, and other types of silent areas. The effects evoked by ventilation noise have also been found to be related to the type of activity being conducted. Annoyance and performance thus not only seemed to be linked to the physical character of exposure, i.e., noise level, frequency characteristics, and length of exposure, but also mental and manual activity, complexity, and monotonony of the work. The effects can be described in terms of annoyance, discomfort, and fatigue, with consequences on performance and increased mental load. The silent areas where ventilation noise may be most frequently experienced are often synonymous with areas and activities most sensitive to the exposure.

8:45
1aAA4. The relevance of low-frequency sound properties for performance and pleasantness. Kerstin Persson Waye and Johanna Bengtsson (Dept. of Environ. Medicine, The Sahlgrenska Acad. of Gothenburg Univ., Box 414, 405 30 Gothenburg, Sweden, kerstin.persson-waye@enmed.gu.se)

The sound environment in the workplace has been found to influence performance, stress, mood, and well-being after work. However few studies can provide dose-response relationships and little is known of the importance of sound-quality aspects for adverse effects on critical tasks or task requirements. We have, during the last 8 years, been engaged in studies investigating the critical performance effects due to the presence of low frequencies (20–200 Hz) in sounds. The main hypotheses on critical effects derived from studies in the general environment were that low-frequency noise induced great annoyance, concentration difficulties, and was difficult to filter out or habituate to. On the other hand, results from truck drivers indicated that low-frequency sounds may lead to reduced alertness and increased sleepiness. In total, three studies were designed with regard to these hypotheses, all of them with the intention to be applicable to office and control room environment, using equivalent A-weighted sound-pressure levels of 40 and 45 dB. The fourth study investigated the importance of sound properties in low-frequency sounds for the perception of pleasantness. The results will be presented and discussed in relation to noise assessment aspects. [Work supported by Swedish Council for Working Life and Social Research.]

9:05
1aAA5. Relating human productivity and annoyance to indoor noise criteria systems. Erica E. Bowden and Lily M. Wang (Architectural Eng. Prog., Univ. of Nebraska–Lincoln, 245 PKI, 1110 S. 67th St., Omaha, NE 68182, ebowden@mail.unomaha.edu)

The goal of this research is to determine a noise criteria system which best relates the effects of background noise to human productivity and annoyance. A number of indoor noise criteria systems are currently used to rate the background noise in built environments, including noise criteria (NC), balanced noise criteria (NCB), room criteria (RC), room criteria Mark II (RC-Mark II), and others. Many questions still remain about the accuracy of these predictors in assessing human response to background noise under the variety of ambient noise situations encountered. To support the use of any individual criterion, subjective testing was performed under a range of background noise situations and statistically related to the various noise criteria predictors listed above. Subjects completed an annoyance survey and performed typing and proofreading tasks in an acoustically controlled environment under 12 simulated background noise settings. These settings varied across three sound levels and four spectral qualities. Subjective testing methodology and results are presented. [Work supported by INCE and ASHRAE.]

9:25

This study investigates the effects of interference speech and the built acoustical environment on human performance, and the possibility of designing spaces to architecturally meet the acoustical goals of office and classroom environments. The effects of room size, geometry, and acoustical parameters on human performance are studied through human subject testing. Three experiments are used to investigate the effects of distracting background speech on short-term memory for verbally presented prose under constrained laboratory conditions. Short-term memory performance is rated within four different acoustical spaces and five background noise levels, as well as a quiet condition. The presentation will cover research methods, results, and possibilities for furthering this research. [Work supported by the Program in Architectural Acoustics, School of Architecture, Rensselaer Polytechnic Institute.]
1aAA7. Environmental impact on workplace performance. Derrick P. Knight, Steven J. Orfield, and Thomas J. Smith (Orfield Labs., 2709 E. 25th St., Minneapolis, MN 55406, derrick@orfieldlabs.com)

Since 1998, Orfield Laboratories has led research efforts into the effects of the office environment on worker performance through the Open Plan Working Group. This collaboration of researchers, designers, and facility managers works with OPWG in gathering environmental data (including noise, lighting, and thermal measurements) and subjective occupant ranking data. The OPWG then employs statistical analysis to correlate the environmental measurements, occupant performance, and occupant preference measurements. Through efficient measurement stations and Web-based surveys, Orfield Laboratories has helped many companies to begin to view their employees as their greatest operational profit center. This presentation will discuss the effect that noise as well as other environmental factors has on work performance.

10:05–10:15 Break

Contributed Papers

1aA9. Classroom acoustics in public schools: A case study. Carmen H20849 Curitiba, Brazil

Classroom noise levels often exceed recommendations and, in large scale retrospective studies, it has been suggested that higher noise levels often correlate significantly with poorer academic performance [e.g., Shield., et al. (2002)]. However, experimental data on the performance of individual children are limited. This study therefore examined the effect of noise on the performance of children in grades 3–4 and 7–8 on standardized tests of oral reading, silent reading, and vocabulary (the Gray Oral Reading Test, the Gray Silent Reading Test, and the Peabody Picture Vocabulary Test). Each child completed parallel forms of a test in quiet and in classroom noise presented at 60 dB SPL. Required speech was presented at +10 S/N. Results from grouped data showed significantly reduced performance in noise only on the silent reading task and only for the older group of children. However, across tasks, when the effects of noise were evaluated as a function of children's quiet performance levels, the noise effect was shown to be significant for the children performing at above average levels in quiet. These findings suggest that the effect of classroom noise may vary significantly across tasks and children. [Work supported by CLLRNet.]

1aAA10. Marking emergency exits and evacuation routes with sound beacons utilizing the precedence effect. Sander J. van Wijngaarden, Adelbert W. Bronkhorst, and Louis C. Boer (TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands)

Sound beacons can be extremely useful during emergency evacuations, especially when vision is obscured by smoke. When exits are marked with suitable sound sources, people can find these using only their capacity for directional hearing. Unfortunately, unless very explicit instructions were given, sound beacons currently commercially available (based on modulated noise) led to disappointing results during an evacuation experiment in a traffic tunnel. Only 19% out of 65 subjects were able to find an exit by ear. A signal designed to be more self-explanatory and less hostile-sounding (alternating chime signal and spoken message “exit here”) increased the success rate to 86%. In a more complex environment—a mock-up of a ship’s interior—routes to the exit were marked using multiple beacons. By applying carefully designed time delays between successive beacons, the direction of the route was marked, utilizing the precedence effect. Out of 34 subjects, 71% correctly followed the evacuation route by ear (compared to 24% for a noise signal as used in commercially available beacons). Even when subjects were forced to make a worst-case left–right decision at a T-junction, between two beacons differing only in arrival of the first wave front, 77% made the right decision.
11:15

1aAA12. Allowable floor-impact sound levels in apartment buildings.
Jin Yong Jeon and Jeong Ho Jeong (School of Architectural Eng., Hanyang Univ., Seoul 133-791, Korea, jyjeon@hanyang.ac.kr)

The purpose of this study is to review the sound-isolation efficiencies of structural materials in multistory residential buildings, in which floor-impact sound was considered as the major source of residents complaints. Floor, walls, and ceiling were selectively treated, and the floor-impact sound was measured using standard noise sources such as the impact ball, the bang machine, and the tapping machine. The noise from the impactors was analyzed and the relationship between the sound levels and the subjective responses was investigated. Results showed that the overall sound level of the impact ball is slightly higher than that of a bang machine, although the impact ball has a lower impact force. It was also found that the noise from the impact ball is similar to the noise of children running and jumping. In addition, when the noise from the impact ball was evaluated in both laboratory and in situ conditions, the allowable noise level was found to be 54 dB (Li, $F_{\text{max}}$, AW).

11:18

1aAA13. Human annoyance and reactions to hotel room specific noises.
Ian L. Everhard (Acoust. Dept., Columbia College Chicago, 33 E. Congress, Chicago, IL 60605)

A new formula is presented where multiple annoyance sources and transmission loss values of any partition are combined to produce a new single number rating of annoyance. The explanation of the formula is based on theoretical psychoacoustics and survey testing used to create variables used to weight the results. An imaginary hotel room is processed through the new formula and is rated based on theoretical survey results that would be taken by guests of the hotel. The new single number rating compares the multiple sources of annoyance to a single imaginary unbiased source where absolute level is the only factor in stimulating a linear rise in annoyance [Fidell et al., J. Acoust. Soc. Am. 66, 1427 (1979); D. M. Jones and D. E. Broadbent, “Human performance and noise,” in Handbook of Noise Control, 3rd ed., edited by C. M. Harris (ASA, New York, 1998), Chap. 24; J. P. Conroy and J. S. Roland, “STC Field Testing and Results,” in Sound and Vibration Magazine, Acoustical Publications, pp. 10–15 (July 2003)].

11:45

Reinhard Weber, Ingo Baumann, Nils Freese, and Volker Mellert (Inst. of Phys., Faculty V, Oldenburg Univ., D-26111 Oldenburg, Germany)

In the frame of the multinational European project HEACE Health effects of aircraft cabin environment [www.heace.org] experiments have been carried out to investigate the effects of different environmental factors in an aircraft on performance, comfort and health of flight and cabin crew. Tests were run in aircraft cabin simulators where temperature, humidity and sound could be adjusted in a controlled manner because only limited possibility exists of systematically changing these factors in-flight. In a multi-factorial $3 \times 3 \times 3$ design these tests simulated real flights with real cabin crew that was hired for the test and passenger. The research on passengers responses was done in cooperation with the European FACE Technology Platform (FACE Friendly aircraft cabin environment). This paper focuses on the effects of noise on the comfort on the cabin crew. It presents unexpected order effects of noise assessments and reports on the dependency of the ratings of noise and of other environmental factors on the assessed comfort. [The investigation is granted by the EU-Commission under HEACE G4RC-CT-2001-00611.]

Monday Morning, 24 May 2004
Bats hunting flying insects adjust the design of their echolocation calls according to the situation in which they forage and stage in an attack. Changes in call design across attack sequences alert other bats within earshot to the presence of prey, demonstrating a continuum in roles for biosonar signals between orientation and communication. Many aerial-feeding bats change the design of their echolocation calls in the presence of echolocating conspecifics. Bats may change frequency parameters, durations, and/or intensities of their calls. While a variety of free-tailed bats (Molossidae Otomops martiensseni, Tadarida teniotis, Molossus molossus) consistently change their echolocation calls when more than one bat is flying in an area, at least one sheath-tailed bat (Emballonuridae Taphozous perforatus) does not. Changes in echolocation calls may maximize jamming avoidance and/or enhance the communicative function of the calls. The data for molossids support the hypothesis that when hunting some species fly in formation. Here, variation in individual call design could provide positional information and reduce the chances of mid-air collisions.

Field studies of four sympatric parrot species in Costa Rica are revealing several possible functions for the well-known ability of parrots to mimic new sounds throughout life. Despite earlier suggestions that this might facilitate exchanges of environmental information, all data so far suggest that vocal mimicry in the wild is associated with mediation of the fission/fusion of groups of parrots and/or of conflicts between mated pairs. Recent results using array recording and interactive playback will be summarized, and several technical problems created by the mechanisms of parrot vocal signal production discussed. [Research supported by NSF Grant IBN-022927 and by continued encouragement and logistics provided by the staff of the Area Conservation Guanacaste (Costa Rica).]

The role of physical and physiological constraints in determining the performance limits on the tempo and frequency bandwidth of birdsong was investigated. One series of experiments examined the mechanism by which a vocal mimic, the northern mockingbird (Mimus polygnotos), copied the songs of other species with which it was tutored as a juvenile. Other experiments analyzed the motor basis of special canary (Serinus canaria) syllables eliciting sexual responses from females. In each case, the mechanism of vocalization was determined by measuring the respiratory dynamics and sound produced on each side of the songbirds duplex vocal organ, the syrinx. When mockingbirds copied the songs of other species the accuracy of their copy depended on the accuracy with which they reproduced the motor pattern used by the tutor species. Motor difficulty of various acoustic features was assessed by the accuracy of its copy. The high repetition rate, broadband canary syllables preferred by females required especially demanding bilateral motor skills. The results indicate that constraints on the rate of respiratory ventilation and bilateral syringeal coordination can set an upper limit on syllable repetition rate and frequency bandwidth. [Work supported by NIH and NSF.]
10:35
1aAB7. Array measurement of echolocation signals on the melon of harbor porpoises (*Phocoena phocoena*). Whitlow Au, Kelly Benoit-Bird (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734, wau@hawaii.edu), Ronald Kastelein (Sea Mammal Res. Co., Harderwijk, The Netherlands), and Ted Cranford (San Diego State Univ., San Diego, CA)

The melon of odontocetes has been hypothesized to be a focusing body that channels echolocation signals produced within the nasal region of the animal’s head into the water. The acoustic field of two echolocating harbor porpoises (*Phocoena phocoena*) was measured at the melon’s surface with an array of four broadband hydrophones embedded in suction cups. The clicks detected by each hydrophone were simultaneously digitized at a sampling rate of 500 kHz and stored on a PC. Digital still photographs of the array were taken before each echolocation trial to measure the hydrophone’s position. The shape and dimensions of the melon were measured with a flexible shape-retaining measuring device. The axis of the echolocation beam was found to be approximately 5.6–6.1 cm from the edge of the animal’s upper lip along the midline of the melon, which coincides with the axis of the low-density lipid core of the melon. Click amplitudes dropped off rapidly (12–14 dB) from the maximum at hydrophones 3.5 cm apart, providing support for the melon-focusing hypothesis. Changes in the waveform on consecutive hydrophones suggest that the porpoises can manipulate either the shape of the melon or the output of the sources within 0.086 s.

10:55
1aAB8. Echolocation click rates and behavior of foraging Hawaiian spinner dolphins. Kelly J. Benoit-Bird and Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734)

Groups of spinner dolphins work together to actively aggregate small animals in the deep-scattering layer that serve as their prey. Detailed information on dolphin foraging behavior, obtained with a 200-kHz multibeam sonar (Simrad MS2000), made it possible to correlate echolocation and foraging. Fifty-six groups of spinner dolphins foraging at night within a midwater micronekton sound-scattering layer were observed with the sonar. During sonar surveys, the rates of whistles and echolocation clicks were measured using four hydrophones at 6-m depth intervals. Significant differences in click rates were found between depths and between the different stages of foraging. Groups of foraging dolphins ranged in size from 16 to 28 dolphins. Click rates were not significantly affected by the number of dolphins in a foraging group. Contrary to initial predictions, click rates were relatively low when sonar data indicated that pairs of dolphins were actively feeding. Highest echolocation rates occurred within the scattering layer, during transitions between foraging stages. Whistles were only detected when dolphins were not in a foraging formation and when animals were surfacing. This suggests clicks may be used directly or indirectly to cue group movement during foraging.

11:15
1aAB9. Stuttering: A novel bullfrog vocalization. Andrea Simmons and Dianne Suggs (Dept. of Psych., Brown Univ., Providence, RI 02912, Andrea_Simmons@brown.edu)

The advertisement call of male bullfrogs (*Rana catesbeiana*) consists of a series of individual croaks, each of which contains multiple harmonics with a missing or attenuated fundamental frequency of approximately 100 Hz. The envelope of individual croaks has typically been represented in the literature as smooth and unmodulated. From an analysis of 5251 advertisement calls from 17 different choruses over two mating seasons, we show that males add an extra modulation (around 4 Hz) to the envelope of individual croaks, following specific rules. We term these extra modulations stutters. Neither single croak calls nor the first croak in multiple croak calls contains stutters. When stuttering begins, it does so with a croak containing a single stutter, and the number of stutters increases linearly (plus or minus 1 stutter, up to 4 stutters) with the number of croaks. This pattern is stable across individual males (*N* = 10). Playback experiments reveal that vocal responses to stuttered and nonstuttered calls vary with proximity to the stimulus. Close males respond with nonstuttered calls, while far males respond with stuttered calls. The data suggest that nonstuttered calls are used for aggressive or territorial purposes, while stuttered calls are used to attract females.

11:35
1aAB10. Impact of the chorus environment on temporal processing of advertisement calls by gray treefrogs. Joshua Schwartz, Kenneth Huth, and Jeffrey Lasker (Dept. of Biol., Pace Univ., 861 Bedford Rd., Pleasantville, NY 10570, jswartz2@pace.edu)

Male gray treefrogs advertise for mates using calls that consist of a series of pulses. Pulse duration, interpulse interval, and pulse shape determine whether a call is recognized as a conspecific signal by females. Females use call rate and call pulse number to assess relative calling performance by males, and prefer males that display high calling efforts. However, within choruses call overlap among males and background noise can compromise the ability of females to detect and correctly interpret temporal information in calls. Phonotaxis tests using calls suffering from different patterns of overlap or with internal gaps were used to investigate specific consequences of interference and masking as well as mechanisms that might alleviate such problems. Our data indicate that females do not employ a process analogous to phonemic restoration to “fill in” missing call segments; however, if a sufficient percent of call elements fall within species-specific ranges, females may ignore call anomalies. Additional findings are generally consistent with those from a recent study on anuran auditory midbrain neurons that count and indicate that inappropriate pulse intervals can reset the pulse counting process. [Work supported by NSF and a Pace University Eugene M. Lang Research Fellowship.]
Session 1aBB

Biomedical Ultrasound/Bioresponse to Vibration and Signal Processing in Acoustics:
High Frequency Imaging

Jeffrey A. Ketterling, Cochair
Riverside Research Institute, 156 William Street, New York, New York 10038-2609

Orlando Aristizábal, Cochair
Skirball Institute of Bimolecular Medicine, 540 First Avenue, New York, New York 10016

Chair’s Introduction—7:55

Invited Papers

8:00

1aBB1. Current and future innovations in high-frequency ultrasonic transducers and arrays. K. Kirk Shung (Dept. of Biomed. Eng., Univ. of Southern California, Los Angeles, CA 90089, kkshung@usc.edu)

High-frequency ultrasonic imaging is considered the next frontier in ultrasonic imaging. Commercial high-frequency scanners often termed ultrasonic backscatter microscope (UBM) all use scanned single element transducers at frequencies higher than 30 MHz with a frame rate lower than 30 frames/s. The engineering of single element transducers and linear arrays at these frequencies has been problematic and may be addressed from the end of developing better piezoelectric materials and from the end of developing novel fabrication methodologies. The further improvement of UBM can be benefited greatly by the availability of superior single element transducers of wider bandwidth and higher sensitivity. Various novel piezoelectric materials, including 2-2 composites, fine grain materials, high or low dielectric constant materials, and fabrication methods that utilize conventional dice and fill, thick film, and MEMS have been explored to fulfill this need. Prototype linear arrays higher than 30 MHz have been developed and tested and preliminary images obtained. These recent developments will be discussed in this talk. [Work supported by NIH.]

8:20


Pulsed lasers can generate ultrasound through thermoelastic expansion of a thin optical absorber. By carefully designing the optical absorbing structure, efficient transduction is possible for a number of biomedical applications including high-frequency imaging, microfluidics, and sensing. The major key for efficient optoacoustic transduction in biomedical applications is to engineer a nearly perfect optical absorber possessing a large coefficient of thermal expansion with acoustic properties well matched to a water medium. We have obtained an optoacoustic efficiency increase of over 20 dB compared to conventional approaches using a thin, optically absorbing layer consisting of polydimethylsiloxane (PDMS) and carbon black spin coated onto a clear PDMS substrate. This structure has been extensively analyzed both experimentally and analytically and seems to provide opportunities for a wide range of optoacoustic devices. In this talk we show how PDMS-based optoacoustic transduction can be used for high-frequency imaging using longitudinal waves and acoustic tweezing using Lamb waves. The basic mechanism of optoacoustic transduction will be described, and specific devices will be presented.

8:40

1aBB3. Ultrasound biomicroscopy in mouse cardiovascular development. Daniel H. Turnbull (Skirball Inst. of Biomolecular Medicine and Depts. of Radiol. & Pathol., New York Univ. School of Medicine, New York, NY 10016, turnbull@saturn.med.nyu.edu)

The mouse is the preferred animal model for studying mammalian cardiovascular development and many human congenital heart diseases. Ultrasound biomicroscopy (UBM), utilizing high-frequency (40–50-MHz) ultrasound, is uniquely capable of providing in vivo, real-time microimaging and Doppler blood velocity measurements in mouse embryos and neonates. UBM analyses of normal and abnormal mouse cardiac function will be described to illustrate the power of this microimaging approach. In particular, real-time UBM images have been used to analyze dimensional changes in the mouse heart from embryonic to neonatal stages. UBM-Doppler has been used recently to examine the precise timing of onset of a functional circulation in early-stage mouse embryos, from the first detectable cardiac contractions. In other experiments, blood velocity waveforms have been analyzed to characterize the functional phenotype of mutant mouse embryos having defects in cardiac valve formation. Finally, UBM has been developed for
real-time, in utero image-guided injection of mouse embryos, enabling cell transplantation and genetic gain-of-function experiments with transfected cells and retroviruses. In summary, UBM provides a unique and powerful approach for in vivo analysis and image-guided manipulation in normal and genetically engineered mice, over a wide range of embryonic to neonatal developmental stages.

9:00


The eye is ideally suited for diagnostic imaging with very high frequency (>35 MHz) ultrasound (VHUF) because of its peripheral location and cystic structure. VHUF allows high resolution visualization of pathologies affecting the anterior segment of the eye, including tumors, cysts, foreign bodies, and corneal pathologies. We developed a series of prototype instruments suitable for ophthalmic studies using both polymer and lithium niobate transducers, with digitization of radiofrequency echo data at up to 500 MHz. While initially using linear scan geometries, we subsequently developed an arc-shaped scan matched to the curvature of the 0.5-mm-thick cornea to circumvent the effect of specular deflection of the ultrasound beam produced by the cornea's curved surface. This technique allowed us to obtain data across the entire cornea and determination of the thickness of each corneal layer, including the epithelium (approximately 50 microns in thickness) and the surgically induced interface produced in LASIK, the most common form of refractive surgery. By scanning in a series of meridians, and applying optimized signal processing strategies (deconvolution, analytic signal envelope determination), corneal pachymetric maps representing the local thickness of each layer can be generated and aid in diagnosis of surgically induced defects or refractive abnormalities.

9:20

1aBB5. Scattering and statistical models for very-high-frequency ultrasonic monitoring of tumor therapy. Frederic L. Lizzi (Riverside Res. Inst., 136 William St., New York, NY 10038)

Our laboratories examined spectral changes in ultrasonic backscatter from tumors undergoing various forms of treatment, with the goal of developing noninvasive treatment monitoring. Human tumor explants in athymic mice were treated in vivo with ultrasonic hyperthermia, high-intensity focused ultrasound, and chemical agents. Pre- and posttreatment spectral examinations were conducted using very-high-frequency ultrasound (40-MHz center frequencies). Physical scatterer properties were estimated from measured spectral data. Spatially averaged spectra showed that successful treatment progressively increased acoustic concentration by about 3 dB. It also produced average changes of 2 μm in scatter sizes of cell-sized (10 μm) structures; these perturbations may be associated with apoptosis, vacuole formation, and frank cellular disruption. These findings are consistent with clinical observations (40 MHz) following radiotherapy of ocular tumors. Theoretical scattering and statistical models were applied to specify transducer and processing schemes that will permit these changes to be mapped in high resolution images, rather than evaluated as spatially averaged values. This requires improved estimator precision for spectral arrays of small scatterers near the Rayleigh size limit. Results showed that this goal can be achieved with 50-MHz annular arrays combined with adaptation of spectral calibration and processing procedures; these are now being implemented for clinical application.

Contributed Papers

9:40

1aBB6. Imaging the shear flow of complex fluids using high-frequency ultrasound. Sébastien Manneville, Lydiane Bécu, Annie Colin (Ctr. de Recherche Paul Pascal—CNRS, Ave. Schweitzer, F-33600 Pessac, France), Mickaël Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, ESPCI, 75005 Paris, France)

Complex fluids show fascinating properties under flow due to the existence of a mesoscopic scale located between the microscopic and the macroscopic scales. For instance, uncoiling of polymer molecules under flow leads to well-documented shear-thinning behaviors. More surprisingly, inhomogeneous flows of visco-elastic fluids are observed in simple shear experiments even at very low Reynolds numbers. Indeed, due to strong coupling between the flow and the fluid microstructure, shear may induce new structural organizations that coexist in the flow leading to shear localization or to shear bands. Whereas classical rheology only yields global data such as the viscosity averaged over the whole sample, the present work is devoted to local ultrasonic velocimetry in sheared complex fluids. A 1D high-frequency (36 MHz) speckle tracking technique is presented that allows the spatio-temporal study of various inhomogeneous unsteady flows. Moreover, preliminary results using a 2D ultrafast flow imaging technique at 12 MHz show that images of the flow can be obtained with an unprecedented temporal resolution.

9:55

1aBB7. Breast tissue characterization with high-frequency scanning acoustic microscopy. R. E. Kumon, I. Bruno, B. Heartwell, and E. Maeva (Dept. of Phys., Univ. of Windsor, 401 Sunset Ave., Windsor, ON N9B 3P4, Canada, kumon@uwindsor.ca)

We have performed imaging of breast tissue using scanning acoustic microscopy (SAM) in the range of 25–50 MHz with the goal of accurately and rapidly determining the structure and composition throughout the volume of the samples. In contrast to traditional histological slides, SAM images can be obtained without special preparation, sometimes even without sectioning, but with sufficiently high spatial resolution to give information comparable to surface optical images. As a result, the use of high-frequency SAM at the time of breast lumpectomy to identify disease-free margins has the potential to reduce reoperative rates, patient anxiety, and local recurrence. However, only limited work has been performed to characterize breast tissue in the frequency range above clinical ultrasound devices. The samples are 4-cm²-thick sections (2–3 mm) taken from mastectomies and preserved in formalin. They are placed between two plates and immersed in water during imaging. Attenuation images are acquired by focusing the acoustic beam at the top and bottom of the samples, although better results were obtained for bottom focusing. For purposes of comparison and identification of histological features, acoustical images will be presented along with optical images obtained from the same samples. [Work supported by CIHR.]
A scanning acoustic microscope (SAM) is described that can measure attenuation, sound speed, impedance, and backscatter coefficient of a sample. The SAM consists of a spherical focused transducer (10–40 MHz) which operates in pulse-echo mode and is scanned in a 2D raster pattern over a sample. A plane wave analysis is presented which allows the impedance, attenuation, and phase velocity of a sample to be recovered from the front and back echoes. Axial attenuation model is used to validate the attenuation, and phase velocity measurements. The analysis of Chen et al. [UFFC, 445, 515] is used to obtain the backscatter coefficient based on the echoes generated by subwavelength scatterers within the sample. The results for the impedance, attenuation and phase velocity were validated for high- and low-density polyethylene against published results. The SAM was used to measure the impedance, attenuation, phase velocity, and backscatter coefficient for the medial and adventitial layers of an in vitro human femoral artery. The theory was extended to account for propagation through multiple layers. The SAM measurements of acoustic properties were in good agreement with previously published results. [Work supported by the NSF through the Center for Subsurface Sensing and Imaging Systems.]

The computation of quantitative ultrasonic parameters such as the attenuation or backscatter coefficient requires compensation for diffraction effects. In this work a simple and accurate diffraction correction method for skin characterization requiring only a single focal zone is developed. The advantage of this method is that the transducer need not be mechanically repositioned to collect data from several focal zones, thereby reducing the time of imaging and preventing motion artifacts. Data were first collected under controlled conditions from skin of volunteers using a high-frequency system (center frequency = 33 MHz, BW = 28 MHz) at 19 focal zones through axial translation. Using these data, mean backscatter power spectra were computed as a function of the distance between the transducer and the tissue, which then served as empirical diffraction correction curves for subsequent data. The method was demonstrated on patients patch-tested for contact dermatitis. The computed attenuation coefficient slope was significantly (p<0.05) lower at the affected site (0.13 ± 0.02 dB/mm/MHz) compared to normal skin (0.2 ± 0.05 dB/mm/MHz). The mean backscatter level was also significantly lower at the affected site (6.7 ± 2.1 in arbitrary units) compared to normal skin (11.3 ± 3.2). These results show diffraction corrected ultrasonic parameters can differentiate normal from affected skin tissues.

An ultrasound biomicroscopy (UBM) system operating between 40 and 50 MHz has been developed for imaging and detecting blood flow in normal and genetically modified mouse embryos, which has contributed to the understanding of the functional consequences of specific genetic defects affecting cardiovascular development. In this scanner, UBM images and Doppler blood velocity waveforms are acquired simultaneously with separate transducers, enabling independent optimization of either the imaging or Doppler transducer. Air backed, 1.5-mm-diameter, unfocused 44-MHz LiNbO3 Doppler transducers designed for the 40- to 50-MHz pulsed Doppler ultrasound system have been developed. The fabrication and characterization of the electromechanical and ultrasonic beam properties of these transducers will be described. The two-way insertion loss at 44 MHz was measured to be −16 dB before tuning, a 60% improvement over the piezo-polymer transducers typically used for UBM-Doppler. Beam-plot data demonstrate the expected collimated near-field pattern to a depth of 13 mm, providing reasonable lateral resolution (1.0–1.5 mm) over a wide depth of field. The utility of these Doppler transducers for imaging small blood vessels such as the dorsal aorta in normal and mutant mouse embryos with defects in cardiac valve formation has been demonstrated.

very-high-frequency ultrasound (VHFU) transducer technology is currently experiencing a great deal of interest. Traditionally, researchers have used single-element transducers which achieve exceptional lateral image resolution although at a very limited depth of field. A 5-ring focused employed in this study makes use of pairs of unfocused reciprocal transducers (send/receive) to provide redundant measurements and simplicity. The method was tested by calibrating a hydrophone for which the manufacturer calibration data were available. Our test data, with error bars of about 1.7 dB, fell well within the range of the manufacturers data, which have error bars of about 1 dB. The results for the passive PVDF probe (Imotec) showed that sensitivity decreased by about one order of magnitude as the temperature decreased from 22 °C (i.e., room temperature) to 3 °C. A small increase is observed as temperature rises from 22 °C to 37 °C although the sensitivity remains about the same order of magnitude.

Nonlinear imaging has already shown improved image resolution compared to fundamental imaging at lower frequencies (2–4 MHz). The required steps to obtain nonlinear images at high frequencies are presented here. The transmit frequency was from 20 to 60 MHz and on receive, pulse-inversion, the first, second, and the third harmonic were used for imaging. Experiments were conducted using a wire phantom and in vitro from a human femoral artery. PVDF transducer (6 mm diameter, 10 mm focal length, central frequency 42 MHz, −6 dB bandwidth 38 MHz) was used with a filtered amplifier (46.0 and 50.6 dBm output). The pulse-inversion experiments at 40 and 60 MHz had lateral resolution (~6 dB contours) 64 and 52 μm and axial 43 and 31 μm. The fundamental experiments at 20, 30, 40, and 60 MHz had lateral resolutions 104, 89, 74, 57 μm and axial 72, 86, 46, 35 μm. The second harmonic at 40 and 60 MHz had lateral resolution 62 and 47 μm and axial 52 and 38 μm. The third harmonic at 60 MHz had lateral resolution 48 μm and axial 42 μm. Nonlinear tissue images had improved resolution compared to fundamental imaging, though the improvement was small for the second harmonic experiment.

An ultrasound biomicroscopy (UBM) system operating between 40 and 50 MHz has been developed for imaging and detecting blood flow in normal and genetically modified mouse embryos, which has contributed to the understanding of the functional consequences of specific genetic defects affecting cardiovascular development. In this scanner, UBM images and Doppler blood velocity waveforms are acquired simultaneously with separate transducers, enabling independent optimization of either the imaging or Doppler transducer. Air backed, 1.5-mm-diameter, unfocused 44-MHz LiNbO3 Doppler transducers designed for the 40- to 50-MHz pulsed Doppler ultrasound system have been developed. The fabrication and characterization of the electromechanical and ultrasonic beam properties of these transducers will be described. The two-way insertion loss at 44 MHz was measured to be −16 dB before tuning, a 60% improvement over the piezo-polymer transducers typically used for UBM-Doppler. Beam-plot data demonstrate the expected collimated near-field pattern to a depth of 13 mm, providing reasonable lateral resolution (1.0–1.5 mm) over a wide depth of field. The utility of these Doppler transducers for imaging small blood vessels such as the dorsal aorta in normal and mutant mouse embryos with defects in cardiac valve formation has been demonstrated.

The computation of quantitative ultrasonic parameters such as the attenuation or backscatter coefficient requires compensation for diffraction effects. In this work a simple and accurate diffraction correction method for skin characterization requiring only a single focal zone is developed. The advantage of this method is that the transducer need not be mechanically repositioned to collect data from several focal zones, thereby reducing the time of imaging and preventing motion artifacts. Data were first collected under controlled conditions from skin of volunteers using a high-frequency system (center frequency = 33 MHz, BW = 28 MHz) at 19 focal zones through axial translation. Using these data, mean backscatter power spectra were computed as a function of the distance between the transducer and the tissue, which then served as empirical diffraction correction curves for subsequent data. The method was demonstrated on patients patch-tested for contact dermatitis. The computed attenuation coefficient slope was significantly (p<0.05) lower at the affected site (0.13 ± 0.02 dB/mm/MHz) compared to normal skin (0.2 ± 0.05 dB/mm/MHz). The mean backscatter level was also significantly lower at the affected site (6.7 ± 2.1 in arbitrary units) compared to normal skin (11.3 ± 3.2). These results show diffraction corrected ultrasonic parameters can differentiate normal from affected skin tissues.

The disadvantage of this method is that the transducer need not be mechanically repositioned to collect data from several focal zones, thereby reducing the time of imaging and preventing motion artifacts. Data were first collected under controlled conditions from skin of volunteers using a high-frequency system (center frequency = 33 MHz, BW = 28 MHz) at 19 focal zones through axial translation. Using these data, mean backscatter power spectra were computed as a function of the distance between the transducer and the tissue, which then served as empirical diffraction correction curves for subsequent data. The method was demonstrated on patients patch-tested for contact dermatitis. The computed attenuation coefficient slope was significantly (p<0.05) lower at the affected site (0.13 ± 0.02 dB/mm/MHz) compared to normal skin (0.2 ± 0.05 dB/mm/MHz). The mean backscatter level was also significantly lower at the affected site (6.7 ± 2.1 in arbitrary units) compared to normal skin (11.3 ± 3.2). These results show diffraction corrected ultrasonic parameters can differentiate normal from affected skin tissues.
annular array, a transducer geometry that permits an increased depth of
field via electronic focusing, has been constructed. The transducer is fab-
cricated with a PVDF membrane and a copper-clad Kapton film with an
annular array pattern. The PVDF is bonded to the Kapton film and pressed
into a spherically curved shape. The back side of the transducer is then
filled with epoxy. One side of the PVDF is metallized with gold, forming
the ground plane of the transducer. The array elements are accessed elec-
trically via copper traces formed on the Kapton film. The annular array
consists of 5 equal-area rings with an outer diameter of 1 cm and a radius
of curvature of 9 mm. A wire reflector target was used to test the imaging
capability of the transducer by acquiring B-scan data for each transmit/
receive pair. A synthetic aperture approach was then used to reconstruct
the image and demonstrate the enhanced depth of field capabilities of the
transducer.

MONDAY MORNING, 24 MAY 2004

VERSAILLES TERRACE, 10:10 A.M. TO 12:00 NOON

Session 1aNS

Noise: Noise Effects and Hearing Protection

Elliott H. Berger, Chair
E-A-R/Aearo Company, 7911 Zionsville Road, Indianapolis, Indiana 46268-1657

Chair’s Introduction—10:10

Contributed Papers

10:15
1aNS1. The development of noise-induced hearing loss in military trades. Sharon M. Abel and Stephanie Jewell (Defence Res. and Development Canada—Toronto, 1133 Sheppard Ave. W., Toronto, ON M3M 3B9, Canada)

An investigation is in progress to determine risk factors for the develop-
ment of noise-induced hearing loss in Canadian Forces personnel. A
total of 1057 individuals representing a wide range of military trades have
contributed their current audiogram, first audiogram on record, and, re-
sponses to a 56-item questionnaire. The protocol for the hearing test was
standardized and conformed to current audiological practice. The items
included in the questionnaire related to demographics, occupational and
nonoccupational noise exposure history, training in and utilization of per-
sonal hearing protectors, and factors other than noise which might affect
hearing (e.g., head injury, ear disease, exposure to solvents, and the use of
medications). Analysis are under way to determine the average current
hearing thresholds as a function of frequency and change relative to base-
line values at recruitment for groups defined by trade, rated noise hazard,
and years of service. Preliminary results suggest ways to improve the
training of personnel with respect to the effects of both occupational and
nonoccupational noise exposure and methods of implementing hearing
conservation strategies. The role of head injury, history of ear disease,
and the use of medications appear to be small. [Work supported by Veterans
Affairs Canada.]

10:30
1aNS2. Estimating the precision error in hearing protector ratings. William J. Murphy, John R. Franks (Hearing Loss Prevention Section, NIOSH, 4676 Columbia Pkwy., MS C-27, Cincinnati, OH 45226, wjm4@cdc.gov), and Peter B. Shaw (NIOSH, Cincinnati, OH 45226)

The variance associated with hearing protector rating methods is de-

erived for the Noise Reduction Rating, Noise Reduction Rating Subject Fit,
and Single Number Rating methods. Real-ear attenuation at threshold
(REAT) data have an error that is a function of stimulus frequency, pro-
tector fit, protector type, and the subject performance during occluded and
uncoccluded threshold estimation. In addition to the error associated with
conducting the REAT tests, there is an additional error that results from
applying the rating to a noise spectrum different from the spectrum used to
derive the rating. For this paper, a distinction is made between the accu-
racity of the rating and the precision of the rating. The precision is an

inherent property of the subject panel, protector type and fit, and the fre-
quencies tested. The accuracy will be attributed to the application of the
rating to an arbitrary noise. [ Portions of work supported by the U. S.
Environmental Protection Agency (EPA) promulgated a labeling regula-
tion specifying a Noise Reduction Rating (NRR) measured in decibels
dB). In the intervening 25 years many concerns have arisen over this
regulation. Currently the EPA is considering proposing a revised rule. This
report examines the relevant issues in order to provide recommendations
for new ratings and a new method of obtaining the test data. The conclu-
sion is that a Noise Reduction Statistic for use with A weighting (NRS),

An important question to ask in regard to hearing protection devices
(HPDs) is how much hearing protection they can provide. With respect to
the law, at least, this question was answered in 1979 when the U. S.
Environmental Protection Agency (EPA) promulgated a labeling regula-
tion specifying a Noise Reduction Rating (NRR) measured in decibels
(dB). In the intervening 25 years many concerns have arisen over this
regulation. Currently the EPA is considering proposing a revised rule. This
report examines the relevant issues in order to provide recommendations
for new ratings and a new method of obtaining the test data. The conclu-
sion is that a Noise Reduction Statistic for use with A weighting (NRS),

an A–A’ rating computed in a manner that considers both intersubject and
interspectrum variation in protection, yields sufficient precision. Two such
statistics ought to be specified on the primary package label—the smaller
one to indicate the protection that is possible for most users to exceed, and

a larger one such that the range between the two numbers conveys to the
user the uncertainty in protection provided. Guidance on how to employ
these numbers, and a suggestion for an additional, more precise, graphi-
cally oriented rating to be provided on a secondary label, are also in-
cluded.

11:00
1aNS4. Real-world performance of hearing protection systems in Army aviation. William A. Ahroon and Martin B. Robinette (U.S. Army Aeromedical Res. Lab., P.O. Box 620577, Fort Rucker, AL 36362-0577, william.ahroon@us.army.mil)

Methods to measure the performance of hearing protective devices
using naive subjects were developed to better estimate the hearing protec-
tion that can be reasonably achieved in operational environments. A ‘’pref-
11:15


Attenuation of acoustic impulses by various models of foam, premolded, formable and semi-insert earplugs was measured. The impulses were generated by a loudspeaker system in the 115–130 range of levels, and by a blast of air expanding from a cylinder in the 150–170 level range. Transmission loss method was used to determine the difference between the peak level under the earplug and outside the earplug. The earplugs were inserted into in-house designed cylindrical or conical couplers attached to an artificial test fixture. The measurements showed that attenuation of high-level impulses depends on the type of coupler and the depth of earplug insertion. Attenuation measured using a conical coupler was consistently about 5 dB lower than for cylindrical coupler. Decreasing the depth of earplug insertion from 75% to 50% was especially critical at high level impulses and lowered attenuation by about 6 dB. For various models of foam and premolded earplugs, attenuation of high-level impulses varies by more than 20 dB. Particularly high attenuation was obtained for formable earplug made of modified wax. [Work supported by the State Committee for Scientific Research Grant No. III-7.03, III-6.07.]

11:30


A database of roughly 10,000 adult persons—from 18 years to 70 years—was analyzed with a procedure that accounts for normal aging of the ear. Auditory performance of good-hearing persons of various groups was determined: office personnel, construction workers, university students, airline pilots, dentists, orchestra musicians, fans of discotheques, avoiders of discotheques, Tibetan nomads, Chinese peasants living without technical noise, etc. Pure-tone auditory threshold—based on pulsed signals—was analyzed from 125 Hz to 10 kHz. The Leq of the long-term acoustic environment was estimated, using acoustic measurements. Results confirm the well-known fact that excessive noise levels damage the ear. However, at lower levels, sound can apparently improve the sense of hearing, as measured by the auditory threshold. Ranking the various groups according to their hearing performance reveals that the best-hearing groups are all living and working in a loud acoustic environment. In these groups aging of the ear is reduced. Groups living in an environment with very low sound levels do not hear well. Apparently the auditory system needs training, in order to develop its full potential, and to keep functioning well.

11:45


A special procedure was developed to separate ears with auditory damage from those without damage. It was applied to a database of roughly 10,000 persons, containing pure-tone audiograms and other information. Persons with medical problems of the ear were excluded. For analysis, groups of persons with similar ways of life were examined: college students, orchestra musicians, dentists, Tibetan nomads, Chinese city dwellers, etc. Looking specifically at ears that suffered noise-induced auditory damage, the damages are different in men and women. At frequencies below about 2 kHz, men typically have less hearing loss than women. However, above 2 kHz, men have more hearing loss than women. This effect is quite strong and highly significant. It could be found in every group that is large enough. It is present in persons going regularly to discotheques, as well as in Tibetan nomads, who live without any technical noise, but use fire crackers now and then. This effect appears not related to the acoustic environment, but can be the result of differences in vulnerability between men and women. It can also be found in youngsters and, to some extent, in children.
Session 1aPAa

Physical Acoustics: Thermoacoustics, Bubbles and Films

David L. Gardner, Chair
Los Alamos National Laboratory, Condensed Matter and Thermal Physics Group, Los Alamos, New Mexico 87545

Contributed Papers

8:00

1aPAa1. DeltaEZ: A visual interpreter for DeltaE. Gordon P. Smith (Western Kentucky Univ., Bowling Green, KY 42101)

Thermoacoustic devices typically begin their development utilizing DeltaE, a numerical simulator developed by researchers at the Los Alamos National Laboratories. The program is highly popular amongst the thermoacoustics community for its ability to successfully describe the physical parameters and geometries that a thermoacoustic device requires to operate. However, the program is text-based, and can be somewhat counterintuitive to the novice user. Alternate, graphically based, programs have been developed for the thermoacoustic community, meeting with mixed reactions. This approach differs from earlier efforts in that no attempt is made to replace DeltaE itself. A front-end interface possessing the intuitive benefits of graphical programs has been created, which then runs DeltaE. Details of the beta version of the program will be presented in the hopes of gathering volunteers for further development.

8:15

1aPAa2. A small-scale, thermoacoustic-Stirling electric generator for deep-space applications. Scott Backhaus (Condensed Matter and Thermal Phys. Group, Los Alamos Natl Lab., Los Alamos, NM 87545), Michael Petach, and Emanuel Tward (Northrop Grumman Space Technol., Redondo Beach, CA 90277)

Although thermoacoustic-Stirling hybrid engines (TASHE) have not been previously coupled to transducers to produce useful electric power, they have demonstrated high thermal-to-acoustic power conversion efficiencies. Electric generation is investigated by coupling a small TASHE to an electrodynamic linear alternator with an emphasis on satisfying NASA’s need for a small, lightweight, efficient electric generator for deep-space missions. The combined goals of low mass and high efficiency require the TASHE to have the largest acoustic power output possible from a minimum enclosed volume, which imposes a relation between various impedances of the TASHE’s lumped-element loop. The design of the TASHE and alternator used in this generator will be reviewed, performance data presented, and possible improvements discussed. [Work supported by NASA.]

8:30

1aPAa3. Characterization of regenerator materials for thermoacoustic refrigeration. Jin Liu and Steven Garrett (Grad. Prog. in Acoust., Penn State, P.O. Box 30, State College, PA 16804)

A bellows bounce test apparatus using a mechanically resonant gas-filled metal bellows, coupled to moving-magnet linear motor, contained within a pressure vessel, has been developed for the testing of porous regenerator materials that might be useful for thermoacoustic refrigeration. The apparatus provides a combination of operating pressures, gases, amplitudes, and frequencies that can cover a large range of Reynolds numbers and Lautrec numbers (ratio of the hydraulic radius to the thermal penetration depth). We will report initial results obtained with this apparatus using two ceramic Celcor samples (400 and 600 cells/in²) over four orders-of-magnitude in Reynolds number. The experimental results were in good agreement with the linear acoustic approximation at low Reynolds numbers, but deviate at higher Reynolds numbers for the viscous component of the pressure drop. For low Lautrec numbers, the dimensionless pressure drop depends on both Lautrec and Reynolds numbers. For Lautrec numbers greater than one (Poiseuille flow regime), there is no frequency (Lautrec) dependence. Similitude analysis collapses all data on both viscous (in-phase) and inertial (quadrature) pressure drops suggesting that introduction of the Lautrec number (or some dimensionless variable related to penetration depth) was required and measurements on samples of more complicated geometries should provide reliable results.

8:45

1aPAa4. Thermoacoustic boundary layers near the liquid—vapor critical point. Keith A. Gillis, Josi L. Shinder, and Michael R. Moldover (Process Measurements Div., NIST, Gaithersburg, MD 20899-8360, keith.gillis@nist.gov)

The sound attenuation in resonators filled with xenon at its critical density \( \rho_c \) was calculated and measured as a function of the reduced temperature \( \tau = (T - T_c)/T_c \) (\( T_c \) is the critical temperature.). Over the temperature and frequency ranges of the measurements \( 10^{-3} < \tau < 10^{-1} \), 0.1 kHz < f < 7.5 kHz, the attenuation was dominated by the thermal boundary layer. The model predicts that the attenuation at the boundary first increases as \( \tau \) decreases and then saturates when the effusivity of the xenon exceeds that of the solid. The effusivity is \( \epsilon = \sqrt{\rho C P \lambda} \), where \( C_p \) is the isobaric specific heat and \( \lambda \) is the thermal conductivity. The model correctly predicts (±1.0%) the quality factors \( Q \) of resonances measured in a steel resonator (\( \rho_c = 6400 \text{ kg/m}^3 \cdot \text{s}^{-1/2} \)), it also predicts the observed increase of the \( Q \) by up to a factor of 8, when the resonator is coated with a polymer (\( \rho_c = 370 \text{ kg/m}^3 \cdot \text{s}^{-1/2} \)). The thickness \( \delta_T \) of the thermal boundary layer in the xenon decreases as \( \tau \) decreases until \( 2\pi\sqrt{\gamma/(\rho c^2)} \approx 1 \). (\( \gamma \) is the bulk viscosity, \( c \) is the speed of sound.) For smaller \( \tau \), \( \delta_T \) is predicted to become complex and increase. [Work supported by NASA.]

9:00


The thermal buffer tube is the component of thermoaoustic systems intended to transmit sound and to provide a thermal barrier between the hot and cold heat exchangers located at the opposite ends of the tube. In high-power operating regimes of thermoaoustic devices, the sound amplitude is high enough to cause significant mass (Rayleigh) streaming in the thermal buffer tube. The streaming flow convects heat, which can degrade the overall performance of a thermoaoustic system by thermally short-circuiting the hot and cold heat exchangers. This leads to increased heat consumption by prime movers and to reduced cooling capacity in refrigerators. The thermal buffer tube is characterized by a significant axial variation of temperature, so effects of gravity on the streaming should be important. A simple analytical model allows quantitative estimation of gravity effects, showing that gravity may significantly suppress the
steering under certain conditions. The important parameters of the system and operating conditions that affect the steering in the presence of gravity are identified. [Work supported by DOE’s Office of Science.]

9:15


As traveling-wave thermoacoustic devices are scaled to higher power, the cross-sectional areas of their regenerators increase proportionally in order to keep pressure losses constant. Unfortunately, this allows the possibility of a spatially dependent acoustic streaming instability internal to their regenerators due to weak thermal and hydrodynamic communication transverse to the acoustic axis. Calculations show one possible result is two regions of opposed streaming forming a recirculating acoustic streaming cell similar to a convection cell. If the streaming is vigorous enough, it can transport enthalpy between heat exchangers at different temperatures, reducing the efficiency of the thermoacoustic device. An experimental apparatus is being used to search for this instability. A summary of the theory and results from the experimental search will be presented. [Work supported by DOE’s Office of Science.]

9:30

1aPAa7. Maximum in the damping of shape oscillations of foam drops near the critical void fraction. Hai Wang, Li Liu, Gregory J. McDaniel, and Glynn R. Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, haiw@bu.edu)

Small spherical samples of aqueous foam ("foam drops") of varying gas volume fraction are acoustically levitated in an ultrasonic field. The normalized natural frequency and damping ratio are determined by treating a foam drop as a damped linear oscillator and measuring shape mode frequency response. The observed natural frequencies are compared to a wet foam model without any fitting parameters. Good agreement is found for gas volume fractions ranging from 0.01 to 0.87. The observed damping ratio of a foam drop exhibits a maximum for gas volume fractions 0.5 to 0.7. The data are inverted to infer the shear elastic modulus and its dependence on gas volume fraction for dry foams. [Work supported by NASA.]

9:45

1aPAa8. Magnetic resonance measurement of dynamics of cavitating fluid. Igor Mastikhin, Benedict Newling, Bruce J. Balcom, and Derrick Green (MRI Ctr., Phys. Dept., Univ. of New Brunswick, 8 Bailey Dr., Fredericton, NB E3B 5A3, Canada, mast@unb.ca)

The prevalent methods used in studies of cavitation are optical and acoustical. They are sensitive to changes in optical and/or acoustical transparency and are not applicable to studies of opaque media. Magnetic resonance (MR) methods, on the other hand, can be applied to arbitrarily opaque media, providing both dynamic and molecular information. A hindrance to implementation of MR methods in cavitation research is their relatively long measurement time: they cannot compete with optics or acoustics when a researcher needs a snapshot technique to study quickly changing processes on a microsecond scale. In this work, we show feasibility of an application of MR to studies of dynamics of cavitating fluid, with measurements of spatially resolved velocity spectra and other above-mentioned parameters. [Work supported by NASA.]

10:00–10:15 Break

10:15

1aPAa9. BUSS: Final results from an experiment to determine the effect of gravity on sonoluminescence. Charles Thomas, Ronald Roy, and R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

The goal of the Boston University Sonoluminescence in Space (BUSS) project was to investigate the dependence of SBSL on ambient acceleration by performing experiments both on the KC-135 and in a land-based lab at Boston University. In this talk the final results of the KC-135 portion of the project will be discussed. Measurements of the maximum bubble radius, bubble position, and peak light intensity emitted (Rmax, z, and Ipeak, respectively) were made during parabolic flight while keeping the acoustic pressure constant. Trends in Rmax, z, and Ipeak are studied by picking out their extrema from the data for each parabola. Additionally, the measured time-dependent Rmax, z, and Ipeak are compared to a quasi-static model which was developed to predict changes in these variables as a function of the changing acceleration and cabin pressure of the KC-135 during flight. Given the relatively good agreement of the model with measured data, it is concluded that the changes in Rmax, z, and Ipeak can be explained simply by considering the change in hydrostatic pressure caused by the variable acceleration. [Work supported by NASA.]

10:30

1aPAa10. Modulated single-bubble sonoluminescence: Dependence of phase of flashes, their intensity and rise/decay times on viscosity, the modulation strength, and frequency. Igor Mastikhin and Borko Djurkovic (MRI Ctr., Phys. Dept., Univ. of New Brunswick, 8 Bailey Dr., Fredericton, NB E3B 5A3, Canada, mast@unb.ca)

The single-bubble sonoluminescence (SBSL) signal was studied for the case of driving frequency modulated by lower frequency with an offset. In our work, the driving frequency of 28 kHz and the modulation frequencies of 25–1000 Hz were used. The modulation strength of 0.2, 0.5, and 0.8 was defined as the difference of highest and lowest pressures over modulation period. The measurements were performed for water–glycerol mixtures of various viscosities. The measured SBSL signal appeared as a train of flashes for modulation frequencies below 250 Hz, and as a continuous modulated signal for higher frequencies. At the same frequency, the flashes covered similar phase intervals for different modulation strengths and, accordingly, pressure ranges. At higher glycerol concentrations (up to 24%) both the intensity and the stability of flashes increased, due to damped shape instabilities and reduced dancing; however, the phase interval of flashes remained about the same. Such phase-locked behavior can be explained by translational movements of the bubble due to modulated Bjerknes force and changes in the symmetry of the bubble collapse. The changes in intensities and rise/decay times can serve as a measure of the gas exchange between the bubble and its surroundings during silent and luminescent intervals.

10:45

1aPAa11. Chemical oscillations in bubbles: Resolving the mystery of chaotic sonoluminescence. Charles Thomas, Ronald Roy, R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215), and Joachim Holzfluss (TU Darmstadt, 64289 Darmstadt, Germany)

One aspect of the Boston University Sonoluminescence in Space project included 1-g experiments to map the position of stable bubbles in (P0, Rmax) space, in water for a range of dissolved gas concentrations. While performing these experiments a regime was found in which Rmax and the position of the bubble oscillate periodically on a slow timescale despite a constant acoustic pressure. Histograms of the phase of light emission from a bubble in such an oscillatory state reveal a broadband distribution. These measurements partially explain previous measurements.
of quasiperiodic and chaotic sonoluminescence. We hypothesize that the oscillations are caused by the competing mass transfer mechanisms of rectified diffusion (growth) and chemical reaction (shrinkage). [Work supported by NASA.]

11:00

1aPAb12. Faraday film patterns. R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215) and John R. Saylor (Clemson Univ., Clemson, SC 29634-0921)

Locally flat liquid interfaces that are periodically forced to vibrate above a threshold amplitude exhibit an instability to waves of finite amplitude and half the forcing frequency. These waves (first observed and explored by Faraday) self-organize into a variety of patterns depending on experimental parameters. We report here the use of Faraday waves in liquid/particulate mixtures to cause deposition of the suspended particles in patterns that mimic the overlaying Faraday wave field. In this work the liquid/particulate mixture is deposited on a solid substrate in the form of a thin film. Hence the wave motion has a significant effect on the velocity field near the solid surface affecting the deposition locations of the particles. Subsequent evaporation of the liquid phase results in a stable, patterned, particulate film. We demonstrate this method experimentally at millimeter scales using relatively large particles, and discuss the possibility of using this method at higher frequencies to create particulate patterns at the micrometer and nanometer scales. Simulations of gold/silver particulate mixtures reveal variations in the particulate deposition pattern for these two particle types due solely to their density difference, suggesting a possible method for generating composite materials tailored on a very fine scale.

11:15

1aPAb13. Modeling of the bulk acoustic wave effects of electrode thin films in piezoelectric resonators. Sonal Srivastava and Yook-Kong Yong (Rutgers Univ., 623 Bowser Rd., Piscataway, NJ 08854)

The modeling of very high frequency piezoelectric resonators such as the thickness shear mode resonators is computationally and memorywise very intensive. The finite element method must be sufficiently fine so as to prevent the mesh impedance from interfering with the wave acoustics of the vibrating modes. The thin film of electrodes introduces additional complexity to the finite element models, hence an efficient method for treating the acoustic effects of such films is needed. Since the films on the resonator surfaces are stress-free, the film stiffness is small and negligible, and only the inertial effects need be modeled. Finite elements models are used to demonstrate the effects of electrode film stiffness and mass on the bulk acoustic wave characteristics of the piezoelectric plate. Results are presented to show the upper limits of the electrode film thickness upon which only the inertial effects are predominant. Results are also shown to demonstrate the efficiency of modeling only the inertial effects and neglecting the stiffness effects of thin electrode films.

MONDAY MORNING, 24 MAY 2004

NEW YORK BALLROOM B, 8:20 TO 11:30 A.M.

Session 1aPAb

Physical Acoustics: Recent Advances in Buried Landmine Detection I

James M. Sabatier, Cochair
University of Mississippi, National Center for Physical Acoustics, Coliseum Drive, University, Mississippi 38677

Bradley Libbey, Cochair
Night Vision and Electronic Sensors Directorate, 10221 Burbeck Road, Fort Belvoir, Virginia 22060

Invited Papers

8:20

1aPAb1. Nonlinear seismo-acoustic landmine detection. Dimitri Donskoy (Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030)

The seismo-acoustic methods are among the most promising emerging techniques for the detection of landmines. Numerous field tests have demonstrated that buried landmines manifest themselves at the surface through linear and nonlinear responses (or signatures) to acoustic/seismic excitation at the frequencies below 1000 Hz. The linear signatures are due to lower impedance of soil above softer mine, especially in the vicinity of mines resonance. The nonlinear signatures are explained by high contact nonlinearity at the mine–soil interface. These phenomena are utilized in high provability/low false alarm detection methods pioneered by University of Mississippi (linear detection) and Stevens Institute of Technology (nonlinear detection). A simple mass-spring model of the mine-soil system [Donskoy et al., J. Acoust. Soc. Am. (2002)] explains and provides an analytical tool for analysis and prediction of both: linear and nonlinear signatures with respect to depths, soil, and mine types. This presentation provides an overview of theoretical and experimental investigations conducted at Stevens over the last 4 years with the emphasis on nonlinear detection techniques. Among major accomplishments are discovery and explanation of mines resonance behavior; soil/depth effect on buried mines resonances; discovery and analysis of nonlinear acoustic interactions at the soil–mine interface; and development of nonlinear quadratic and intermodulation detection algorithms based on dual-frequency excitation.
1aPAb2. Nonlinear acoustic experiments involving landmine detection: A connection between mesoscopic/nanoscale effects in geomaterials. Murray S. Korman (Dept. of Phys., U.S. Naval Acad., Annapolis, MD 21402)

The vibration interaction between the top-plate interface of a buried plastic landmine and the soil above it appears to exhibit many characteristics of the mesoscopic/nanoscale nonlinear effects that are observed in geomaterials like rocks (sandstone) or granular materials. Experiments are performed with an inert VS 1.6 anti-tank mine that is buried 3.6 cm deep in dry sifted loess soil. Airborne sound at two primary frequencies \( f_1 = 120 \text{ Hz} \) and \( f_2 = 130 \text{ Hz} \) undergo acoustic-to-seismic coupling. Interactions with the compliant mine and soil generate combination frequencies that, through scattering, can affect the vibration velocity at the surface. Profiles of the soil surface particle velocity at \( f_1 \) and \( f_2 \) and the nonlinearly generated \( f_1 - f_2 \) component are characterized by a single peak. Doubly peaked profiles at \( 2f_1 + f_2 \) and \( 2f_2 + f_1 \) are attributed to the familiar mode shape of a timpani drum. Near resonance, the bending (a softening) of a family of tuning curves for the soil surface vibration over a landmine exhibits a linear relationship between the peak frequency and the corresponding peak particle velocity, which also exhibit hysteresis effects. [Work supported by U.S. Army Communications-Electronics Command RDEC, NVESD, Fort Belvoir, VA.]

9:00

1aPAb3. Landmine detection using acoustic to seismic coupling: Theory and modeling. Roger Waxler (NCPA, Univ. of Mississippi, University, MS 38677, rwax@olemiss.edu) and Doru Velea (Planning Systems Inc., Reston, VA 20191)

It has been demonstrated that buried landmines can be found by insonifying the ground and then measuring the resulting vibration of the ground’s surface. Over a buried landmine there is a large enhancement of the ground’s vibration due to the excitation of mechanical modes of the mine. The resulting ground vibration can be understood as the near-field response to the resonant scattering of sound by the buried landmine. An effective fluid model has been developed. The results and limitations of this model will be discussed.

9:20

1aPAb4. Combining magnitude and phase features in acoustic landmine detection. Tsaipei Wang, James M. Keller (Dept. of Elec. and Computer Eng., Univ. of Missouri–Columbia, Columbia, MO 65211), and Paul D. Gader (Univ. of Florida, Gainesville, FL 32611)

The utility of acoustic-to-seismic coupling systems for landmine detection has been clearly established. They have been shown to be able to detect very low metal content landmines that are difficult to detect for ground-penetrating radars. This technique measures the difference in ground vibration velocity of regions with and without buried landmines when subject to acoustic excitation. For most applications, only the magnitude of the surface velocity is used to construct recognition algorithms. We recently introduced phase-based features in the classification scheme, significantly reducing false alarms at given detection probabilities. Here the focus is on the analysis of ground velocity data collected in the time domain with a moving array of laser Doppler vibrometers. The processing techniques used to extract the magnitude and phase information are described, as well as examples to demonstrate how combined magnitude and phase features help improve the detection of buried mines with weak signatures and reduce false alarms. Simple mass-and-spring models are shown to be useful in understanding the observed features. Also analyzed are the phase signatures of a number of ground regions with buried man-made clutter objects and how they can be useful in separating these clutters from actual landmine regions.

9:40

1aPAb5. Laser-induced acoustic landmine detection. Charles A. DiMarzio, Tianchen Shi, Florian J. Blonigen, and Stephen W. McKnight (Ctr. for Subsurface Sensing and Imaging Systems, Northeastern Univ., Boston, MA 02115)

Laser-induced acoustic (LIA) imaging is a new approach for underground object detection, especially for shallow buried landmine detection. This noncontact detection technique is based on a pulsed laser initiating photo-acoustic interaction in the ground. If sufficient photon energy couples into the ground, generating acoustic signals larger than seismic noise, a high enough mismatch of acoustic impedance between background media and target could provide a clear acoustic image revealing both the target’s position and the 3-D shape information to reduce false alarms. In this paper, we present a model for laser-induced acoustic wave generation, including models of photo-acoustic sources and propagation of sound waves in media, to obtain a better understanding of basic physics behind the process. A photon-induced thermoelastic source and a photon-induced plasma source are investigated. Also, an investigation was performed in terms of attenuation, angular distribution, and phase velocities, and experimental results are reported and compared with theoretical analysis.

10:00–10:15 Break
10:15

1aPA6. Seismic sonar for landmine detection and confirmation.
Thomas G. Muir (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS, tmuir@olemiss.edu), Maneli Zakharia, Aurore Grill, and Emanuel Gestat de Garambe (French Naval Acad., Brest Armes, France)

Impulsive vibration of the ground can generate seismic interface waves of the Rayleigh type. They decay exponentially with depth into the soil, and spread cylindrically with lateral range. At useful frequencies around 100 Hz, they typically travel at speeds around 100 m/s, with wavelengths around a meter. Rayleigh waves can be made to propagate in sonar-like pulses to buried targets, reflect, and return to the sonar for reception and signal processing, providing range, bearing, and information as to target type. We have conducted new experiments and analyses with seismic sonar in a clay soil. A focused array of ten sources and eight receivers (triaxial seismometers) were deployed at a range of 4.5 m to examine a 20-lb. landmine as well as a clump of rocks, and other false targets. After vector polarization processing, the amplitude of the mine target echo was 28 dB above the environmental backscatter. Mine-like target confirmation was provided by Wigner–Ville transformation, which allows separation of manmade from natural targets, by providing unique and identifiable time–frequency–amplitude signatures for each. The potential for target detection as well as a level of target type classification, at relatively long ranges, was demonstrated. [Work supported by the U.S. Army Night Vision Laboratory.]

10:30

1aPA7. Demonstration of an end-fire array Rayleigh wave source for a seismic-acoustic sonar.
Steven R. Baker and Steven E. Rumph (Phys. Dept., Naval Postgrad. School, Code PHuBa, Monterey, CA 93943, srbaker@nps.navy.mil)

A linear array of four vertical-motion sources was deployed on the sand in the near-surf zone of Del Monte Beach, Monterey, CA. The sources were spaced 25 cm apart (approximately one-quarter wavelength at 100 Hz, the nominal operating frequency) and were driven with a transient signal in a sequential fashion so as to preferentially radiate Rayleigh waves in one end-fire direction. Beam patterns were measured at a radius of 3.5 m. Measurements were made with the array directing the radiation toward, away from, and parallel to the surf line. In general, results were in fair agreement with simple (nondispersive) theory, except for the depth of the nulls. A measured front-to-back radiation suppression of approximately 15 dB was routinely achievable.

10:45

1aPA8. Land mine detection by time reversal acousto-seismic method.
Alexander Sutin (Artann Labs., Inc. and Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030), Armen Sarvazyan (Artann Labs., Inc., Trenton, NJ), Paul Johnson, and James TenCate (Los Alamos Natl. Lab., Los Alamos, NM)

We present a concept and results of a pilot study on land mine detection based on the use of time reversal acoustics (TRA). TRA provides a possibility of highly effective concentrating of seismic wave energy in time and space in complex heterogeneous media. TRA focusing of seismic waves on a land mine increases the detection abilities of conventional linear and nonlinear acousto-seismic methods. Such factors as medium inhomogeneities, presence of reflecting boundaries, which could critically limit conventional acoustic approaches, do not affect TRA based method. The TRA mine detection system comprises several air borne or seismic sources and a noncontact (laser vibrometer) device for remote measurements of the surface vibration. The TRA system focuses a seismic wave at a surface point where the vibration is measured. The focusing point is scanned across the search area. The amplitude and frequency dependence of the signal from the seismic wave focusing point and nonlinear acoustic effects are analyzed to assess probability of the mine presence. Preliminary experiments confirmed high focusing ability of the TRA seismo-acoustic system in complex conditions (a laboratory tank with sand) and demonstrated a significant increase in the surface vibration in the presence of mine imitator. [Work supported by DoD grant.]

11:00

1aPA9. Soil effect on landmine vibrations.
Andrei Zagrai, Alexander Ekimov, and Dmitrii Donskoy (Davidson Lab., Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030, azagrai@stevens.edu)

Field tests of the seismo-acoustic landmine detection techniques revealed strong dependence of mines’ vibration signatures on the burial depth, soil physical properties, and environmental conditions. This study discusses the mines’ linear and nonlinear signatures as a function of burial depth, soil type, and moisture content. Experimental results show that the resonance frequency of soil–mine system initially decreases with burial depth and then anomalously increases at greater depths. Soil moisture further amplifies this anomaly. Soil type and moisture content influence mines’ nonlinear signatures as well. It was found that the quadric nonlinear (Q) and cubic nonlinear (C) intermodulation mine vibrations respond differently to the presence of moisture in soil. While Q response diminished in wet soil (after rain), the C response was still highly evident and used for mine visualization. These observations revealed complex behavior of soil–mine system, which is not completely understood or explained. [Work supported by ONR.]

11:15

1aPA10. Effects of elasticity and porosity in modeling of the acoustic-to-seismic transfer function.
Margarita S. Fokina, Vladimir N. Fokin, and James M. Sabatier (Natl. Ctr. for Physical Acoust., 1 Coliseum Dr., University, MS 38655, mfok@olemiss.edu)

Modeling sound interaction with the ground is important both for remote sensing techniques and the elimination of false alarms in the landmine detection application. Elastic and porous-elastic models of the ground are most frequently used for these studies. Though both of these models are well known, few comparisons have been made between the acoustic-to-seismic transfer functions (TF) calculated from these two models for the frequency range 100–1000 Hz, which is typical for geoacoustic applications. In this work, the matrix technique was exploited to solve boundary equations for porous-elastic layers. This technique allows one to obtain the TF and refraction indexes for arbitrary porous-elastic stratifications of ground layers. The results of test computations of original codes are presented. Effects connected with the slow wave in the porous-elastic model are analyzed. Effective parameters of the visco-elastic model are analyzed to give the best fit to frequency dependence of the TF calculated in the frame of porous-elastic layered model. Comparison between real and effective parameters demonstrates the possible accuracy for obtaining properties of porous-elastic ground in the frame of elastic models. [Work supported by ONR Grant N00014-02-1-0878.]
Session 1aPPa

Psychological and Physiological Acoustics: Poster Session I (Poster Session)

Barbara Shinn-Cunningham, Chair

Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, Massachusetts 02215

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.

1aPPa1. Dissonance perception by listeners with sensorineural hearing loss. Jennifer B. Tufts, Marjorie R. Leek, and Michelle R. Molis (Army Audiol. and Speech Ctr., Walter Reed Army Medical Ctr., 6900 Georgia Ave. NW, Washington, DC 20307, jennifer.tufts@na.amedd.army.mil)

The perceived dissonance of two simultaneous tones (a dyad) depends upon their frequency separation (in the case of two pure tones) or their fundamental frequency ratio (in the case of two harmonic complex tones). The purpose of this study was to determine whether the perceived dissonance of puretone and harmonic complex dyads is altered in the presence of sensorineural hearing (SNHL), and, if so, whether this can be explained by the reduced frequency selectivity typically associated with SNHL. Four normal-hearing and four hearing-impaired listeners evaluated the dissonance of puretone and complex harmonic dyads centered at 500 and 2000 Hz. Frequency selectivity was estimated at 500 and 2000 Hz for each listener. People with SNHL rated the dissonance of puretone dyads similarly to the normal-hearing listeners, although auditory sensitivity and frequency selectivity differed considerably between the groups. However, their ratings of harmonic complex dyads did not show the pronounced differences in dissonance as a function of fundamental frequency ratio that were observed in the normal-hearing group. The poorer frequency selectivity of these listeners may have allowed more extensive interactions to occur among harmonic components, resulting in a less clear separation of dissonance and consonance. [Work supported by NIH-NIDCD.]

1aPPa2. Channel interaction in cochlear implants as a function of phase duration and pulse rate. Deniz Baykent and Robert V. Shannon (House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057)

Multichannel cochlear implants stimulate auditory nerves at different locations along the cochlea. It is widely assumed that good speech recognition requires independent activation of distinct tonotopic regions with no interactions between electrodes; interactions between electrodes can produce a result similar to spectral smearing. The electrical pulses produced by multiple electrodes can interact both spatially and temporally. However, little is known about the effect of basic electrical stimulation parameters on the degree of channel interaction. It is possible that high stimulation pulse rates could improve the temporal representation of the signal but reduce the spatial (tonotopic) resolution due to increased electrode interaction. Differences in electrode stimulation configuration (monopolar or bipolar), pulse phase duration, pulse amplitude, and stimulation rate could change the level of electrode interaction. The present study used forward masking to measure electrode interaction patterns as a function of the stimulating pulse phase duration and the stimulation rate. A masker was placed on one electrode and the threshold elevation of a following signal was measured as a function of cochlear location. Forward-masked electrode interaction patterns will be presented for the different stimulation configurations and parameters for several masker levels. [Work supported by NIDCD Grant R01-DC-01526.]

1aPPa3. Place-pitch and excitation patterns in cochlear implant listeners. John J. Galvin III and Qian-Jie Fu (House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057)

In cochlear implants (CIs), an electrode’s pitch is largely determined by the excitation pattern produced by stimulation. For CI users, is pitch perception most strongly influenced by the peak, edge or some intermediate location within the excitation pattern? By varying the stimulation mode at four evenly spaced electrode locations, four spectral profiles were hypothesized: single-peaked, apically weighted, basally weighted or multi-peaked (with an intermediate pitch). Forward-masked excitation patterns and pitch judgments were obtained for the four electrode locations, for all experimental stimulation modes, relative to a standard set of BP1 electrodes. Results showed individual differences in CI users’ excitation patterns and pitch judgments produced by the different stimulation modes. Most subjects’ pitch judgments were sensitive to single, sharp peaks in the excitation pattern; however, subjects differed in their sensitivity to broader, multi-peaked patterns. For widely spaced electrode configurations, some subjects consistently judged pitch according to the apical edge of stimulation, some to the basal edge and others to an intermediate location. For some electrode locations, varying the stimulation mode produced significantly different pitches and excitation patterns. Future speech processing strategies may wish to combine stimulation modes to improve the spectral resolution available with a fixed number of implanted electrodes.

1aPPa4. Spectral profile discrimination ability using cochlear implants: Effects of number of active electrodes and pulse rate. Ward R. Drennan and Bryan E. Pflingst (Kresge Hearing Res. Inst., 1301 E. Ann St., Ann Arbor, MI 48109)

The ability of cochlear implant users to discriminate a change in spectral shape was investigated in listeners with Nucleus CI24M and CI24R(CS) implants. A primary independent variable was the number of active electrodes. Listeners were asked to detect a current increment to one of 3, 7, 11, and 21 active electrodes. Intensity discrimination on one electrode was also evaluated. Stimulation was achieved using a monopolar configuration with 75 μs phase biphasic pulses and a 24 μs interphase gap presented at 250 pps per channel. The pulses were swept in rapid succession from the base to the apex such that the overall rate increased with the number of active electrodes. Sensitivity to differences in spectral shape decreased with increasing number of electrodes. Next, the pulse rate was reduced to 159 pps per channel for the 11 and 21 active electrode conditions. Sensitivity to spectral shape at the slower pulse rate was better than...
that at the higher rate given the same number of active electrodes. The improvement in sensitivity with the slower rate suggests that forward masking of one pulse over a successive pulse diminishes listener’s ability to hear an intensity increment in the spectral profile.

LaPPa5. Psychophysical and speech results from the penetrating auditory brainstem implant (PABI). Robert V. Shannon, Mark E. Robert, and Steve Otto (House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057)

The penetrating auditory brainstem implant (PABI) is a prosthetic device that delivers sound information to the brain via microstimulation of the cochlear nucleus. The design of the speech processor is similar to that in a cochlear implant. Surface electrodes are placed in the lateral recess of the IV ventricle, adjacent to the cochlear nucleus. Penetrating microelectrodes are inserted into the ventral cochlear nucleus. Threshold measures in two patients confirm intra-nuclear placement. Stimulation of the penetrating microelectrodes produced a full range of loudness sensations as well as a wide range of pitch sensations. Initial psychophysical measures show excellent temporal resolution (forward masking, gap detection, modulation detection) on the penetrating electrodes in one patient and poor temporal performance in another. Speech understanding was moderate, with one patient receiving primarily a supplement to lip-reading, while the other patient received modest open-set speech understanding on the third day following initial stimulation. Ongoing studies will characterize the interference between and within surface and penetrating electrode systems and the ability of the patient to integrate speech information across the two types of electrodes. [Work supported by NIDCD Contract N01-DC-00-11.]

LaPPa6. Amplitude modulation detection with cochlear implants: Effects of electrode separation and stimulus level. Anastasios Sarampalis and Monita Chatterjee (House Ear Inst., DAIP, 2100 W. Third St., Los Angeles, CA 90057)

Amplitude modulation (AM) detection performance has been studied in the past with normal-hearing and hearing-impaired populations. The temporal modulation transfer function (TMTF) is a plot of AM detection performance as a function of modulation rate and provides a way of characterizing temporal sensitivity. Typically the TMTF takes the form of a low-pass filter, with performance declining above 50–70-Hz modulation rate. TMTFs have also been measured with cochlear implant patients, showing a similar low-pass characteristic, with a cutoff around 140-Hz rate, while sensitivity to AM was found to increase with increasing current level. The present study investigated the effects of stimulation level and electrode separation on TMTFs with cochlear implant patients. TMTFs were measured for narrow through wide electrode separations and three different (loudness-balanced) percentages of the dynamic range. Preliminary results indicate that sensitivity increases (lower thresholds) with increasing stimulation level, for a given electrode separation. However, comparing TMTFs across different electrode separations, sensitivity is independent of current level, but increases as a function of percentage of dynamic range. In summary, it appears that AM detection performance with cochlear implants depends primarily on sensation level, rather than current level or electrode separation. [Work supported by NIH, R01 DC00397.]


This research is concerned with measurements of temporal onset-order discrimination through the tactual sensory system. Sinusoidal signals were delivered through a multi-finger stimulating device that operates over a frequency range of roughly 0 to 300 Hz and an amplitude range of roughly 0 to 50 dB SL. Measurements were obtained using a one-interval two-alternative forced-choice procedure in which each interval consisted of the random-order presentation of two different sinusoidal signals whose amplitude and duration were roved independently from trial to trial. Thresholds were estimated from psychometric functions of d’ as a function of stimulus-onset asynchrony. Performance was studied as a function of the frequency separation between the two sinusoids presented in a given run for presentation across two different fingers and for presentation to the same finger. Conditions were also included to examine the effects of redundancy of frequency and site of stimulation on performance. Results will be compared across conditions and will be discussed with regard to their implications for the design of tactual displays of speech for persons with profound hearing impairment. [Research supported by Grant No. 5 R01 DC00126 from NIDCD, NIH.]

LaPPa8. Monaural informational masking release in children and adults. Emily Buss, Joseph W. Hall III, and John H. Grose (Univ. of North Carolina at Chapel Hill, 130 Mason Farm Rd., CB7070, 1115 Bioinformatics Bldg., Chapel Hill, NC 27599, ebuss@med.unc.edu)

Informational masking refers to an elevation in signal threshold due to stimulus uncertainty, rather than to energetic masking. This study assessed informational masking and utilization of cues to reduce that masking in children aged 5–9 and adults. We used a manipulation introduced by Kidd et al. [J. Acoust. Soc. Am. 95, 3475–3480 (1994)] in which the signal was a train of eight consecutive tone bursts, each at 1 kHz and 60 ms in duration. Maskers were comprised of a pair of synchronous tone-burst trains whose frequencies were selected from the range spanning 0.2–5 kHz, with a protected region 851–1175 Hz. In the reference condition, where informational masking is pronounced, these maskers were eight bursts and had a fixed frequency within each interval, with new frequencies chosen randomly prior to each interval. Two conditions of masking release were tested: random frequency selection for each masker burst and a masker leading fringe of two additional 60-ms bursts. Both children and adults showed a significant informational masking effect, with children showing a larger effect. Both groups also showed significant release from masking, though initial results suggest that this may have been reduced in the youngest children. [Work supported by NIH, R01 DC00937.]

LaPPa9. Differential effectiveness of cues in informational masking studies. Virginia M. Richards, Ronong Huang (Dept. of Psych., Univ. of Pennsylvania, Ste. 302C, 3401 Walnut St., Philadelphia, PA 19104, richards@psych.upenn.edu), and Gerald Kidd, Jr. (Boston Univ., Boston, MA 02215)

For the detection of a tone added to random multitone maskers, playing a preview of the masker before detection trials can reduce thresholds compared to when there is no preview. In contrast, playing a preview of the signal-plus-masker does not provide a release from masking. This differential effectiveness of cues was examined in several conditions. Using the method of constant stimuli and a yes/no task, observers detected a 1000-Hz tone added to six-tone maskers. Prior to each trial, the frequencies of the masker components were randomly drawn. Two types of cues were tested, either a copy of the masker or a copy of the signal-plus-masker. The cues were presented either before or after the yes/no presentation interval. Finally, data were collected either blocked for each condition or the trials from the four conditions were interleaved. D-prime values were higher when the conditions were blocked than when they were interleaved. The pattern of results was the same in both situations; d’ was highest for pretrial masker cues, lowest for pretrial signal-plus-masker cues, and intermediate when either cue followed the trial interval. [Work supported by NIH/NIDCD.]
Can the processes underlying performance on basic masking tasks be modified with practice? To address this question, 17 listeners completed pre- and posttests in which detection thresholds for a 20-ms 1-kHz tonal signal were measured in three notch-noise simultaneous-masking conditions that differed in the masker duration and relative onset times of the signal and masker, and three forward-masking conditions that differed in the tonal frequencies of the maskers. For the ~8 days between tests, eight of these listeners practiced 240 trials/day on each simultaneous-masking condition. Learning differed across conditions. When the signal and a 20-ms masker began and ended together, both trained and untrained listeners improved, but trained listeners improved more, learning gradually. When the signal began 400 ms after a 500-ms masker, only the trained group improved, learning quickly. When the signal and the 500-ms masker began together, both groups improved equally. Both groups also improved equally on each of the three forward-masking conditions. These results suggest that processes underlying performance on simultaneous- and forward-masking tasks are modifiable, training affects different mechanisms in simultaneous-masking conditions with different temporal characteristics, and the simultaneous-masking mechanisms affected by multi-hour training are not involved in forward masking. [Work supported by NIH.]

**1aPPa11. Training experienced hearing-aid users to identify syllable-initial consonants in quiet and noise.** James D. Miller, Jonathan M. Dalby, Charles S. Watson, and Deborah F. Burleson (Commun. Sci. and Disord. Technol., Inc., Indiana Univ. Res. Park, 501 N. Morton St., Bloomington, IN 47404, jdmiiller@artsic.wustl.edu)

Five experienced hearing-aid users with sensorineural hearing loss were given 14 h of intensive training identifying consonants in quiet and noise. Their performance was compared to that of five similar hearing-aid users with no special training. All listeners had moderate to severe hearing losses and had worn hearing aids for at least 1 year. All were pretrained with a set of 20 consonants combined with three vowels /a, i, u/ as spoken by six different talkers. Pretests were conducted in quiet and in noise (multitalker babble) at moderate signal-to-noise ratios (SNRs). Training was conducted with eight target consonants (TCs). The TCs were in each listener’s middle range of difficulty, and the three most common confusors for each target were individually selected, forming target sets of four consonants. Training was conducted in quiet and noise. During training, trial-by-trial feedback was given and, following an error, the listener could rapidly compare the intended syllable with its confusor. In noise, the SNR adapted to a criterion of 80 correct. There were no differences between training and control listeners on the pretests. After training, there was a significant 5% advantage for the trained listeners. Training generalized to talkers never heard during training. [Work supported by NIDCD.]
4 dB), decision rule (targets = 70.7%, 75%, and 79.4% correct) and initial starting level on thresholds obtained under conditions of inattention and complete attention. Simulations of two-alternative forced-choice (AFC) and three-AFC procedures demonstrate that the best general strategy for measuring thresholds under conditions of inattention is to employ a low percentage (e.g., 70.7%) as a target on the psychometric function combined with a large step size. When a track for an inattentive listener is begun far above threshold, we found that using a large step size for the first two turnarounds (reversals in stimulus level direction) and then switching to a step size that is the initial step size was less effective than using the large step size for the entire threshold estimation run. Behavioral data obtained in a tonal detection-in-noise task with a 2-AFC procedure support the validity of the simulations.

1aPPa16. Effects of N-acetylcysteine on noise-induced temporary threshold shift and temporary emission shift. Martin Robinette (U.S. Army Aeromedical Res. Lab., 6901 Andrews Ave., Fort Rucker, AL 36362-0577, martin.robinette@us.army.mil)

Animal research has shown that antioxidants can provide significant protection to the cochlea from traumatic noise exposure with some benefit when given after the exposure. Similar results in humans would have a significant impact on both prevention and treatment of noise-induced hearing loss. The current study evaluates the effectiveness of N-acetylcysteine (NAC) on temporary threshold shift (TTS) by using both behavioral and physiological measures. Sixteen healthy, normal-hearing subjects were given NAC or a placebo prior to exposure to a 10-min, 102-dB narrowband noise, centered at 2 kHz. This exposure was designed to induce a 10–15-dB TTS. Following the noise exposure, pure-tone thresholds (Bekesy) and transient-evoked otoacoustic emissions (TEOAE) were measured for 60 min to monitor the effects of NAC on TTS recovery. Postexposure measures were compared to baseline data. [Work supported by American BioHealth Group.]

1aPPa17. Acute effects of nicotine on processing of complex stimuli in smokers and nonsmokers. Ashley Harkrider and Mark Hedrick (Dept. of Audiol. and Speech Pathol., Univ. of Tennessee, 457 South Stadium Hall, Knoxville, TN 37996, aharkrid@utk.edu)

Effects of nicotine in the auditory system of normal-hearing smokers and nonsmokers were investigated both behaviorally and physiologically. Discrimination of consonant–vowel speech in quiet and noise was assessed in the presence and absence of a transdermal nicotine patch by measuring categorical boundaries and mismatch negativity (MMN). Data indicate that the effects of nicotine on both behavioral and physiological measures increased with an increase in severity of nicotine-induced symptoms. Smokers showed improved CV discrimination in quiet and noise with nicotine. Additional, smokers exhibited more measurable and significantly sharper boundaries as well as larger MMN areas than nonsmokers in quiet and noise for both placebo and nicotine sessions. MMN data acquired for both quiet and noise, and behavioral data acquired in quiet, indicate that smokers show the greatest improvements in discrimination during nicotine exposure, followed by symptomatic nonsmokers. Asymptomatic nonsmokers show little improvement with nicotine and, on occasion, show decrements in performance. These data may contribute to our understanding of the role of nAChRs in the auditory system, the neural mechanisms that underlie the recognition of sound in quiet and noise, and mechanisms mediating improved information processing and enhanced cognitive performance that serve as reinforcement for continued tobacco use by smokers.

1aPPa18. Fractal EEG analysis with Higuchi’s algorithm of low-frequency noise exposition on humans. Ryszard Panuszka, Zbigniew Damijan, and Cezary Kasprzak (Univ. of Sci. and Technol., Structural Acoust. and Biomed. Eng. Lab., Al. Mickiewicza 30, 30-059 Krakow, Poland, damijan@imir.agh.edu.pl)

Authors used methods based on fractal analysis of EEG signal to assess the influence of low-frequency sound field on the human brain electro-potentials. The relations between LFN (low-frequency noise) and change in fractal dimension EEG signal were measured with stimulations tones. Three types of LFN stimuli were presented; each specified dominant frequency and sound-pressure levels (7 Hz at 120 dB, 18 Hz at 120 dB, and 40 Hz at 110 dB). Standard EEG signal was recorded before, during, and after subject’s exposure for 35 min. LFN. Applied to the analysis fractal dimension of EEG-signal Higuchi’s algorithm. Experiments show LFN influence on complexity of EEG-signal with calculated Higuchi’s algorithm. Observed increase of mean value of Higuchi’s fractal dimension during exposition to LFN.

1aPPa19. Low-frequency sounds and psychological tests at 7, 18, and 40 Hz. Zbigniew Damijan, Cezary Kasprzak, and Ryszard Panuszka (Univ. of Sci. and Technol., Structural Acoust. and Biomed. Eng. Lab., Al. Mickiewicza 30, 30-059 Krakow, Poland, damijan@imir.agh.edu.pl)

Research included the results of tests aimed to determine how LFN (low-frequency noise) with the dominating frequency (7 Hz at 120 dB, 18 Hz at 120 dB, and 40 Hz at 110 dB) influences human brain potentials and understanding the dependency of results achieved in psychological questionnaires. The psychological questionnaires (EPQ-R Eysencks and SSS-5 Zuckermans) were analyzed. Presented issue existences difference, relative influence LFN on human biopotentials, dependence from acquired results in ranges EPQ-R Eysencks and SSS-5 Zuckermans. The test included 96 experiments. Standard EEG potentials, ECG potentials, and EDQ (dermal) were recorded before, during and after subject’s 35-min exposures to LFN. Evident differences in changed bio-signals especially in EEG subject’s dispersion, were easily determined and correlated to questionnaire reports.

1aPPa20. Sound fields in biosphere of the mountain streams and their influence on the human EEG. Cezary Kasprzak, Zbigniew Damijan, and Ryszard Panuszka (Univ. of Sci. and Technol., Structural Acoust. and Biomed. Lab., Al. Mickiewicza 30, 30-059 Krakow, Poland, cekasp@agh.edu.pl)

Low-frequency acoustics fields in the biosphere of the headwater regions in Poland (Bieszczady Mountains and Tatra Mountains) were researched and the influence of observed low-frequency acoustics fields on changes of human EEG signals was analyzed. Several places were found with specific distribution of sound fields where low frequencies dominated. Interesting parameters of distribution of the riverbeds and physical characterization of flows were determined. Standard EEG signal was recorded from subjects until exposure for 20 min. It was ascertained that spectra of sound-pressure levels were with specific shape, very similar for a group of rivers. The influences of SPL on level of power spectrum of EEG signal were analyzed. Changes in the brain waves were observed with increases in delta and decreases in alpha. Tested subjects reported behavior specific for relaxation.
Session 1aPPb

Psychological and Physiological Acoustics: Pitch

Andrew J. Oxenham, Chair
Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Contributed Papers

8:45
1aPPb1. Effects of stimulus level and harmonic resolvability on pitch discrimination. Joshua G. W. Bernstein and Andrew J. Oxenham (MIT Res. Lab. of Electron. and Harvard-MIT Speech & Hearing Biosci. & Technol. Prog., 77 Massachusetts Ave., Cambridge, MA 02139, jgbern@mit.edu

Fundamental frequency (F0) discrimination performance depends largely on harmonic number: only complexes containing harmonics below 10th yield small F0 difference limens (DLs). This may be because only the low harmonics are resolved by the auditory periphery, Auditory filter bandwidths are known to increase with increasing stimulus level, providing a tool for testing the effect of reduced harmonic resolvability on F0 discrimination. F0 DLs were measured for equal-amplitude, random-phase harmonic complexes, with eight different F0s between 75 and 400 Hz, bandpass filtered between 1.5 and 3.5 kHz. Complexes were presented at a constant average sensation level (12.5 dB/component before filtering) in threshold equalizing noise (TEN) with levels of 10, 40, and 65 dB SPL per equivalent rectangular bandwidth (ERBs). If good F0 discrimination requires the presence of resolved harmonics, the transition from poor to good F0 discrimination with increasing F0 should shift to higher F0s at higher levels, because fewer harmonics will be resolved. Preliminary results indicate that three out of four subjects demonstrated this shift at the highest level tested, providing some support for the idea that peripherally resolved harmonics are required for good pitch discrimination performance. [Work supported by NIH Grants R01DC05216 and 5T32DC00038.]

9:00
1aPPb2. Detection and F0 discrimination of harmonic complex tones in the presence of concurrent complexes. Christophe Micheyl and Andrew Oxenham (Res. Lab. of Electron., MIT, Bldg. 36-758, Cambridge, MA 02139-4307, cmicheyl@mit.edu)

This study measured the level of target harmonic complex tones, relative to simultaneous concurrent harmonic or inharmonic maskers, necessary for (a) target detection and (b) discrimination of the target F0. Depending on the condition, the F0 of the target was around 100, 200, or 400 Hz and the average F0 of the masker was either equal to this or 7 semitones lower or higher. Target and masker bands were bandpass filtered between 1200 and 3600 Hz (48-dB/octave slopes). Low-pass noise masked distortion products. Masked thresholds in the target-detection task were generally higher when the average F0 of the masker was lower than that of the target, and were also higher for inharmonic than for harmonic maskers. In the F0 discrimination task, average F0 differences between the target and masker tones consistently improved masked thresholds at the highest F0 used (400 Hz) but not the lowest (100 Hz). Temporal asynchronies between the target and masker and reduced masker-F0 uncertainty had similar effects. These results suggest that peripheral frequency resolution may limit the ability to benefit from auditory segregation cues for both the detection and discrimination of harmonic complex tones embedded in other tones. [Work supported by NIDCD R01DC05216.]

9:30
1aPPb4. Perceiving a change in a nonperceived pitch. Laurent Demaily and Christophe Ramos (CNRS and Univ. Victor Segalen, 146 rue Leo Saignat, F-33076 Bordeaux, France)

Listeners were presented with 300-ms “chords” of five synchronous pure tones, followed after a 0.5–8-s silence by a single pure tone. The frequencies of each chord’s components were selected randomly, but spaced by intervals of between 6 and 10 semitones. In one condition (“up/down”), the single tone following a chord was 1 semitone higher or lower than one of the chord’s three intermediate components; on each trial, the corresponding component was selected randomly and the task was to indicate the direction in which its pitch changed. In another condition (“present/absent”), the single tone following a chord was either identical to one of the three intermediate components or halfway in frequency between two components; the task was to indicate if the single tone was present in the chord or not. Performance was much better in the up/down condition than in the present/absent condition, even though the opposite trend was predictable for an ideal “analytic” listener. Ten listeners reported that, in the up/down condition, they could often perceive the appropriate pitch change without having heard out the relevant component of the chord. These results provide strong evidence for the existence of pitch change detectors in the auditory system.

9:15
1aPPb3. Dominance region for pitch: Effect of duration. Hedwig Gockel, Robert P. Carlyon (MRC Cognition and Brain Sci. Unit, 15 Chaucer Rd., Cambridge CB2 2EF UK, hedwig.gockel@mrc-cbu.cam.ac.uk), and Christopher J. Plack (Univ. of Essex, Colchester CO4 3SQ, UK)

Frequency discrimination thresholds for pure tones increase with decreasing duration, more so at low than at high frequencies. The present study investigated the possible influence of this frequency-dependent effect on the dominance region for pitch (DRP) of complex tones (CTs). The DRP was determined for CTs containing the first seven harmonics with a fundamental frequency (F0) of 250 Hz, and a level of 65 dB SPL per component. The tones had a duration of 16 or 200 ms, including 8-ms ramps. Listeners were presented with a CT with one of the harmonics mistuned upwards or downwards by 3%, followed 500 ms later by a perfectly harmonic CT of the same duration. Listeners had to adjust the F0 of the harmonic tone so that its pitch matched that of the mistuned complex. For all five listeners, the DRP was higher in harmonic number for the short than for the long duration tones. The results are consistent with the idea that pitch is derived from a weighted sum of information from individual harmonics, the individual weights being dependent on duration and related to the discriminability of the harmonics. [Work supported by EPSRC Grant GR/R65794/01.]
10:45

1aPPb8. Low pitches of frequency-transposed stimuli: Salience estimates from a full population-interval model. Peter Cariani (Eaton Peabody Lab., Mass. Eye and Ear Infirmary, 243 Charles St., Boston, MA 02114 cariani@epl.meei.harvard.edu)

Transposed stimuli consisting of half-wave rectified modulators of high frequency carriers allow comparisons of pitch perception across frequency regions. While listeners can discriminate low F0-pitches of simple transposed stimuli (1 Fm-Fc), most cannot detect low pitches of complex transposed stimuli [Fm=300, 400, 500 Hz, Fc=4, 6.35, 10.08 kHz; Oxenham et al., PNAS (2004)]. Auditory nerve (AN) responses to a low-frequency harmonic complex (F0 = 100 Hz, n = 3–5) and its transposition were simulated using filtering (middle ear, 48 gammatone filters, 144 filters, human CF distribution, 100 Hz to 18 kHz), rectification, adaptive gain control, spontaneous activity (three classes), 65 dB SPL simulated level. Pitch salience was estimated by analyzing population-interval histograms using subharmonic interval sieves. The harmonic complex produced high salience (2.05), well above assumed pitch detection threshold (~1.3). Simple transposed stimuli produced lower saliences (1–1.37). Some were below threshold, inconsistent with psychophysical pitch discrimination results. The complex transposed stimulus produced an F0-salience (1.31) near threshold, qualitatively consistent with their mixed detection results. Dramatic differences are thus seen between low-frequency and transposed stimuli. Clearly, for these and other high-frequency stimuli, full AN population-interval models can produce salience estimates that differ substantially from simple intuitions and/or partial implementations. Interpretational caution is therefore counseled.

11:00


Although enormous progress has been made in the development of cochlear implants, the speech reception of implant users is still not comparable to that of normal-hearing listeners. The difference in performance is especially pronounced in complex auditory situations, such as in the presence of competing talkers. The poorer speech reception performance of implant users in complex environments may, at least in part, be due to the poorer fundamental frequency (F0) representation. The present study examined the effects of implant-like processing (i.e., noise-excited vocoding) on F0 discriminability and utility in normal-hearing listeners. F0 difference limens (F0DLs) were measured as a function of the number of vocoder channels. In addition, vowel identification was measured as a function of F0 difference between competing vowels in a double-vowel paradigm, with number of vocoder channels as a parameter. Our findings show that despite the reasonable F0DLs (less than one semitone) with 24- and 8-channel vocoder processing, listeners were unable to benefit from F0 differences between the competing vowels in a double-vowel paradigm. The implications of the findings for pitch theories and cochlear implant design will be discussed. [Work supported by NIDCD Grant R01 DC05216.]
IaSC1. Evaluating the effects of bilingual dominance and language mode on overall degree of foreign accent. Satomi Imai, James E. Flege (Div. of Speech and Hearing Sci., Univ. of Alabama at Birmingham, Birmingham, AL 35294), and Ian R. A. MacKay (Univ. of Ottawa, Ottawa, ON, Canada)

This study evaluated the influence of bilingual dominance and language mode on overall degree of perceived foreign accent. Three groups of Italian–English bilinguals (n = 12 each) were selected according to their ratio of self-rated English/Italian proficiency: English-dominant, balanced, or Italian-dominant. Language mode was manipulated by having participants repeat English phrases before and after similar Italian phrases (E1, E2) and then intermixed with Italian phrases (E3). Native English (NE) listeners rated four English phrases spoken by the bilinguals and 12 age-matched NE controls using a scale that ranged from 1 (strong foreign accent) to 9 (no foreign accent). We hypothesized that if switching into the native language (here, Italian) adversely affects pronunciation of the second language (English), the third repetitions of the English phrases (E3) should be more strongly foreign-accented than the first repetitions (E1). The foreign accent ratings decreased significantly in the following order: NE > English-dominant > balanced > Italian-dominant. That is, all three bilingual groups had detectable foreign accents, and strength of accent depended on bilingual dominance. The language mode effect was significant only for one of the four phrases examined, perhaps because it (mozzarella cheese) has distinctly different phonetic renditions in English and Italian. [Work supported by NIH.]

IaSC2. A cross-language study on perception of Taiwanese stops by non-native listeners. Yueh-chin Chang (Natl. Tsing Hua Univ., 101, Sec. II, Kuang Fu Rd., Hsinchu, Taiwan), Catherine T. Best (Wesleyan Univ., Middletown, CT 06459), and Pierre A. Halle (CNRS-Paris V, 92774 Boulogne-Billancourt, France)

Few reports exist on perception of three-way stop voicing distinctions by non-native listeners whose languages have two-way distinctions that vary phonologically and phonetically. We examined perception of Taiwanese stops by French, Mandarin, and American English listeners. Taiwanese has three voicing categories: unaspirated voiceless (U), aspirated voiceless (A) and voiced (V). Phonologically, English and French have voiced-voiceless contrasts, so these listeners should have similar difficulties with Taiwanese /U-A/. Phonetically, however, English and Mandarin distinguish voiceless unaspirated versus aspirated stops, in which case these listeners should have equivalent difficulties with Taiwanese /U-V/. American and Mandarin listeners’ discrimination supported the second prediction. French listeners discriminated /U-A/ better than /U-V/ for velar stops, but the opposite for labials, possibly because French velar stop VOTs preceding high vowels are longer than Mandarin and American ones. Mandarin listeners discriminated better than French listeners overall. American listeners’ discrimination was poorest. The discrimination results are consistent with identification patterns for the three groups. The findings suggest that speakers of languages with phonologically unaspirated versus aspirated contrasts (Mandarin) can distinguish three voicing types more easily than speakers of languages with a voiced versus voiceless contrast (French, English), especially when the phonetic realizations differ from the phonological distinction (English).

IaSC3. Production and perception of a temporal contrast by native and non-native speakers. Tessa Bent, Ann R. Bradlow (Dept. of Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208, t-bent@northwestern.edu), and Bruce Smith (Univ. of Utah, Salt Lake City, UT 84112)

In previous work, we found that both native and Chinese-accented talkers of English varied considerably in the extent to which they realized the duration contrast for vowels before voiced versus before voiceless consonants. The current study explored the perceptual consequences of this individual variability in production. Twenty native and 35 non-native listeners participated in a minimal pair identification task using stimuli from two native and two non-native talkers who differed substantially in the extent to which they realized the temporal contrast. Stimuli were consonant-vowel-consonant words that differed only in the voicing of the final consonant. Both native and non-native listeners exhibited sensitivity to variation in the production of this temporal contrast: greater identification accuracy was observed for relatively long duration differences. There was also a talker-listener interaction such that non-native listeners identified words produced by the non-native talkers more accurately than those produced by the native talkers. In a follow-up experiment, we investigated the perception-production relationship for this temporal contrast in the non-native listeners. Generally, non-native listeners who were more sensitive to this duration contrast in native-accented English showed a greater duration contrast in English production than non-native listeners who were less sensitive to this contrast. [Work supported by NIH-NIDCD.]

IaSC4. Acoustic influence in crosslanguage phone mapping: Sibilant fricative place mismatch from English to Korean. Yunju Suh (Dept. of Linguist., Stony Brook Univ., Stony Brook, NY 11794-4376, yunju.suh@sunysb.edu)

The loanword adaptation and second language production of the English voiceless palatoalveolar fricative into Korean shows an interesting mismatch in tongue position. Korean palatalizes its alveolar fricative to alveolo-palatal before high front vowels. However, English voiceless palatoalveolar fricative before a high front vowel is adapted as an alveolar...
fricative with the secondary articulation of lip rounding, instead of an alveolar palatal. This paper argues that the failure to use the articulatorily correct Korean alveolar palatal fricative is due to Korean listeners' interpreting English acoustic patterns in terms of the phonetic, especially acoustic, expectations of their native language. That is, Korean listeners attend more to the peak frequency of the English fricative than to the actual tongue position, and map it to their native fricative sound with the spectral peak at the closest frequency. Data collected from female speakers of American English and Korean show that, from fricative midpoint to end, the highest intensity spectral peaks are located at similar frequencies for English palatal-alveolar and Korean rounded alveolar fricatives, while those of the Korean alveolar palatal are 1500–2000 Hz higher. This serves as another piece of evidence that second language loanword adaptation is crucially affected by fine details of L1 and L2 phonetics.

1aSC5. Listening to a non-native speaker: Adaptation and generalization. Constance M. Clarke (Dept. of Psych., Univ. of Arizona, Tucson, AZ 85721)

Non-native speech can cause perceptual difficulty for the native listener, but experience can moderate this difficulty. This study explored the perceptual benefit of a brief (approximately 1 min) exposure to foreign-accented speech using a cross-modal word matching paradigm. Processing speed was tracked by recording reaction times (RTs) to visual probe words following English sentences produced by a Spanish-accented speaker. In experiment 1, RTs decreased significantly over 16 accented utterances and by the end were equal to RTs to a native voice. In experiment 2, adaptation to one Spanish-accented voice improved perceptual efficiency for a new Spanish-accented voice, indicating that abstract properties of accented speech are learned during adaptation. The control group in Experiment 2 also adapted to the accented voice during the test block, suggesting adaptation can occur within two to four sentences. The results emphasize the flexibility of the human speech processing system and the need for a mechanism to explain this adaptation in models of spoken word recognition. [Research supported by an NSF Graduate Research Fellowship and the University of Arizona Cognitive Science Program.]

1aSC6. Language-specific relevance of formant transitions for fricative. Anita Wagner and Marjaim Ernestus (Max-Planck-Inst. for Psycholinguist., Postbus 310, 6500 Nijmegen, The Netherlands, anita.wagner@mpi.nl)

Although the consonant inventories of Dutch, German, English, and Spanish are similar in size, the fricative inventories differ: English and Spanish distinguish labio-dental versus dental fricatives, whereas Dutch and German do not. Three phoneme-monitoring experiments investigated whether the relevance of formant transitions varies across languages, and whether it depends on the types of fricatives in a language. Native Dutch, German, English, and Spanish listeners detected a target fricative, /s/ or /f/, whether it depends on the types of fricatives in a language. Native Dutch, German, and English listeners were unaffected by the misleading formant transitions, in nonwords. Half of the nonwords were cross spliced to produce misleading German listeners were unaffected by the misleading formant transitions, in nonwords. Half of the nonwords were cross spliced to produce misleading

1aSC7. Pseudo-homophony in non-native listening. Anne Cutler (Max Planck Inst. for Psycholinguist., P.O. Box 310, 6500 AH Nijmegen, The Netherlands) and Takashi Otake (Dokkyo Univ., Soka, Japan)

Pseudo-homophony may result when non-native listeners cannot distinguish phonemic contrasts. Thus Dutch listeners have difficulty distinguishing the vowels of English cattle versus kettle, because this contrast is subsumed by a single Dutch vowel category; in consequence, both words may be activated whenever either is heard. A lexical decision study in English explored this phenomenon by testing for repetition priming. The materials contained among 340 items 18 pairs such as cattle/kettle, i.e., contrasting only in those vowels, and 18 pairs contrasting only in /t/ (e.g., right/left). These materials, spoken by a native American English speaker, were presented to fluent non-native speakers of English, 48 Dutch Nijmegen University students, and 48 Japanese Dokkyo University students; the listeners performed lexical decision on each spoken item, and response time was measured. Dutch listeners responded significantly faster to one member of a cattle/kettle pair after having heard the other member earlier in the list (compared with having heard a control word), suggesting that both words had been activated whichever had been heard. Japanese listeners, however, showed no such priming for cattle/kettle words, but did show repetition priming across /t/ pairs such as right/light. Non-native listeners' phonemic discrimination difficulties thus generate pseudo-homophony.

1aSC8. Issues in the measurement of perceptual assimilation. James Harnsberger and Ratree Wayland (Univ. of Florida–Gainesville, Gainesville, FL 32611, jharns@ufl.edu)

This study examined the effect of methodological variables on the fit between predicted discrimination scores based on identification data and actual discrimination data in cross-language speech perception experiments. Such variables include (1) single versus multiple talkers in discrimination test trials; (2) different discrimination test types (e.g., AX, AXB, oddity); and (3) identification tests in which stimuli are presented individually versus stimuli being presented in the same context as they appear in discrimination tests. The optimal pair of identification and discrimination tests, yielding the best match between predicted and actual discrimination scores, can be used in subsequent studies examining perceptual category structure. These methodological variables were examined by presenting American English speakers with two Hindi contrasts, [b]–[p] and breathy voiced dental-retroflex, both in initial position and in [i], [a], or [u] context. The stimuli appeared in a range of categorial discrimination and identification tests. Early results examining the third variable listed above demonstrate that identification tests that present stimuli in the same context as they appear in corresponding discrimination test trials correlate more strongly with discrimination scores (r = 0.72; p < 0.05) than identification tests that present stimuli in isolation (r = 0.58; p < 0.05).

1aSC9. Effects of speaking rate on the perception of phonemic length contrast in Japanese. Hiroaki Kato (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan, kato@atr.jp) and Keiichi Tajima (Housei Univ., Tokyo 102-8160, Japan)

Segment length is distinctive in Japanese, for example, /kaite/ (buyer) versus /kaite/ (seated). Such length contrasts are not necessarily categorically for non-native speakers. To study this property precisely, a series of perception experiments was conducted. A professionally trained native Japanese speaker produced the nonsense word /terete/ at slow, normal, and fast rates with or without a carrier sentence. Either the second vowel or second consonant of each word was gradually lengthened until reaching its

Listeners relied on the transitions even though no dental fricatives occurred in these stimuli, and independently of the speaker's native language.
longer counterpart, i.e., /ete:te/ or /eret:el/, in all rate and carrier conditions using STRAIGHT, a high-fidelity speech analysis, synthesis, and manipulation system [Kawahara et al., Speech Commun. 27, 187–207 (1999)], resulting in 12 stimulus continua. Seven native-Japanese listeners participated in a single-stimulus, two-alternative forced-choice identification task with the method of constant stimuli. The speaking rate of the presented stimuli within a session was either fixed or randomized trial by trial. Results suggest that native listeners’ identification boundaries systematically altered due to changes in speaking rate, whereas their boundaries became unstable in the randomized-rate condition, especially for non-carrier stimuli. These results will be discussed from the viewpoint of second-language phoneme perception and acquisition through comparisons with results from non-native listeners. [Work supported by TAO, Japan.]

1aSC10. Spoken word recognition in English by Japanese listeners: A case of Japanese-accented and unaccented English words. Kiyoko Yoneyama (Dept. of English, Daito Bunka Univ., 1-9-1 Takashimadaira, Itabashi, Tokyo, Japan. yoneyama@ic.daito.ac.jp)

The effect of acoustic mismatch between a listener’s phonological representations and speech input on spoken word recognition was examined with Japanese learners of English. Imai, Plege, and Walley (2003) recently found that native-Spanish listeners showed a larger neighborhood density effect for unaccented English words than Spanish-accented English words, whereas native-English listeners showed a larger neighborhood density effect for Spanish-accented than unaccented words. We hypothesized that phonological mismatches would occur when native-Japanese listeners respond to unaccented English words than Japanese-accented English words. Further, the effect of the mismatch would be expected to be greater for words from dense versus sparse neighborhoods because Yoneyama (2002) found that Japanese listeners showed neighborhood density effect when long counterparts, i.e., /ete:te/ or /eret:el/, etc. were used as word materials, and presented in three conditions: visual-only (V), audio-only (A), and audio-visual (AV) presentations. Identification accuracy of those words produced by two talkers was also assessed. During pretest, the accuracy for A stimuli was lowest, implying that insufficient translation ability and listening ability interact with each other when aurally presented word has to be translated. However, there was no difference in accuracy between V and AV stimuli, suggesting that participants translate the words depending on visual information only. The effect of translation training using AV stimuli did not transfer to identification ability, showing that additional audio information during translation does not help improve speech perception. Further examination is necessary to determine the effective L2 training method. [Work supported by TAO, Japan.]

1aSC11. Effects of acoustic and semantic contexts when learning to identify L2 phonemes in words and sentences. Yuko Ikuma (ATR Human Information Sci. Lab., Kobe Univ., 2-2-2 Hikaridai Keihanni Sci. City, Kyoto 619-0288, Japan, yikuma@atr.jp) and Reiko Akahane-Yamada (Kobe Univ., Kyoto 619-0288, Japan)

Laboratory training experiment was conducted in order to examine the effect of acoustic and semantic contexts when learning second language phoneme perception. Fifty minimal pairs of English words contrasting in /l/ and /r/ were produced by native speakers of American English in three conditions; in isolation (WD), within semantically neutral carrier sentences (NS), and within semantically contextual carrier sentences (CS). Participants were native speakers of Japanese, and were divided into three groups; each was trained to identify /l/ and /r/ in one of above three conditions. In pretest, identification accuracy varied by condition in the order, NS<WD<CS, which replicated the previous study [Rothwell and Akahane-Yamada, J. Acoust. Soc. Am. 112, 2386 (2002)]. It was also shown that the group trained with CS stimuli improved the ability to identify CS stimuli from pretest to post-test, but not WD and NS stimuli. In contrast, the effect of training using WD or NS stimuli generalized to all the stimulus conditions. These results suggest that the perception training utilizing the auditory input, in which acoustic information is the only clue to identify phoneme, is effective in cultivation of aural comprehension. Implications for foreign language education will be discussed. [Work supported by TAO, Japan.]


Komaki and Akahane-Yamada (Proc. ICA2004) used 2AFC translation task in vocabulary training, in which the target word is presented visually in orthographic form of one language, and the appropriate meaning in another language has to be chosen between two choices. Present paper examined the effect of audio-visual presentation of target word when native speakers of Japanese learn to translate English words into Japanese. Pairs of English words contrasted in several phonemic distinctions (e.g., /l–r/, /l–l/, /v–v/, etc.) were used as word materials, and presented in three conditions: visual-only (V), audio-only (A), and audio-visual (AV) presentations. Identification accuracy of those words produced by two talkers was also assessed. During pretest, the accuracy for A stimuli was lowest, implying that insufficient translation ability and listening ability interact with each other when aurally presented word has to be translated. However, there was no difference in accuracy between V and AV stimuli, suggesting that participants translate the words depending on visual information only. The effect of translation training using AV stimuli did not transfer to identification ability, showing that additional audio information during translation does not help improve speech perception. Further examination is necessary to determine the effective L2 training method. [Work supported by TAO, Japan.]

1aSC13. Temporal patterns of native Mandarin Chinese speakers’ productions of English stop-vowel syllable. Yue Wang (Dept. of Linguist., Simon Fraser Univ., 8888 University Dr., Burnaby, BC V5A 1S6, Canada) and Dawn M. Behne (Norwegian Univ. of Sci. and Technol., Trondheim, Norway)

Second language (L2) production can be a kind of interlanguage, a relatively stable system bearing the nature of both the native language (L1) and L2. Within such a system sound components of a syllable may bear their own interlanguage characteristics and yet interact with the other component sounds. The present study investigates temporal patterns of L1–L2 interaction at the syllable level. Audio recordings were made of English stop-vowel syllables produced by native speakers of Mandarin who were fluent in English (ChE). Native English productions (AmE) of these syllables and native productions of Mandarin (ChM) stop-vowel syllables were acquired as native norms. Temporal measures included stop closure duration, voice-onset time (VOT), vowel duration, and syllable duration. Results show that the internal timing components of ChE often deviate from AmE, with the closure duration, VOT, and vowel duration being intermediate to AmE and ChM. However, at the syllable level, ChE productions tend to follow the overall patterns of AmE. Temporal deviations were often compensated by temporal compensation of other components in the syllable, maintaining a balanced consonant/vowel distribution. These findings have implications for a broader understanding of L2 productions.

1aSC14. Intelligibility of non-native Lombard speech for non-native listeners. Chi-Nin Li (Dept. of Linguist., Simon Fraser Univ., Burnaby, BC V5A 1S6, Canada)

Previous study [C-N. Li, J. Acoust. Soc. Am. 114, 2364 (2003)] has shown that foreign-accented Lombard speech is more intelligible than normal speech when presented in noise to native English listeners. This research extends that work and examines the intelligibility of non-native English speakers’ Lombard speech perceived by listeners from the same L1 background. Twelve Cantonese speakers and a comparison group of English speakers read 48 simple true and false English sentences in quiet and in 70 dB of cafeteria noise. Normal and Lombard sentences were masked with noise at a constant signal-to-noise ratio, and presented along with native-accented stimuli to eight native Cantonese speakers who assessed intelligibility by transcribing the sentences in standard English orthography. Analyses indicated that for both groups of speakers, sentences pre-
sented in noise were less well perceived than those presented without noise. The Cantonese speakers’ utterances were more intelligible than were the native English productions. However, in noisy conditions, the Lombard speech of the Cantonese speakers was correctly transcribed less often than their normal utterances, and the English speakers’ Lombard speech was not more intelligible than their normal speech.


One important source of information listeners use to segment speech into discrete words is allophonic variation at word junctures. Previous research has shown that non-native speakers impose their native-language phonetic norms on their second language; as a consequence, non-native speech may exhibit altered patterns of allophonic variation at word junctures. We investigated the perceptual consequences of this for word segmentation by presenting native-English listeners with English word pairs produced either by six native-English speakers or six highly fluent, native-French speakers of English. The target word pairs had contrasting word juncture involving voiceless stop consonants (e.g., why pink/wipe ink; gray ties/great eyes; we cash/weak ash). The task was to identify randomized instances of each individual target word pair (as well as control pairs) by selecting one of four possible choices (e.g., why pink, wipe ink, why ink, wipe pink). Overall, listeners were more accurate in identifying target word pairs produced by the native-English speakers than by the non-native English speakers. These findings suggest that one contribution to the processing cost associated with listening to non-native speech may be the presence of altered allophonic information important for word segmentation. [Work supported by NIH/NIDCD.]

LaSC16. Compensation for phonological assimilation can be triggered by nonspeech sounds. Holger Mitterer (Max-Planck-Institut für Psycholinguistik, P.O. Box 310, 6500 AH Nijmegen, The Netherlands)

Several previous experiments have shown that phonological assimilations are compensated for perceptually by a context-sensitive mechanism. For instance, the Hungarian assimilated form “bar” of the word “bal” (“left”) is recognized as such only if it occurs in a context that allows assimilation (i.e., “balrod” assimilated to “barrol,” “from the left”), but not in other contexts (“barroval”). This “compensation for assimilation” is independent of language experience: Similar results were found with Hungarian and Dutch listeners [Mitterer et al., Proceedings of the 15th International Congress of Phonetic Sciences (2003)]. This suggests that an auditory mechanism underlies compensation. Nonspeech analogs of the phonetic contexts “nal” and “rol” should therefore trigger similar context effects as speech sounds: Three pairs of nonspeech analogs were played after syllables from a “bal”-to-“bar” continuum: First, a broadband noise was used with-analog to the trill in “rol”—or without-analog to the nasal in “nal”—amplitude modulation (AM) at the onset. Second, the same AM manipulation was applied to a 400-Hz tone. Third, frequency modulation replaced the AM of the 400-Hz tone. In each case, the nonspeech sounds had the same effect on identification and discrimination performance as the speech sounds.

LaSC17. The role of low-frequency hearing in speakers sensitivity to delayed auditory feedback. Dragan Barac-Cikjoja and Cara Johnson (Gallaudet Univ., HMB N205F, 800 Florida Ave. NE, Washington, DC 20002)

Sensitivity to delayed auditory feedback (DAF) during speech production was assessed in six normal-hearing, fluent speakers using a two-interval, forced-choice (2IFC) adaptive procedure. Subjects repeated a syllable (either PA, TA, or KA) for approximately 2 s in two successive intervals and listened through headphones for a delay in the speech feedback. Subjects were required to identify which of the two utterances was delayed relative to the production onset. The length of the delay was changed stepwise depending on the accuracy of the subjects response. The estimates of the minimal delay yielding a 71% correct performance were obtained under three experimental conditions. The conditions varied in how the delay affected the low frequency (300 Hz) band (LFB) of the speech feedback signal: (a) LFB was delayed along with the rest of the speech spectrum; (b) LFB was isolated from the rest of the spectrum by band-pass filtering, remained present in the feedback, but was never subjected to a delay; and (c) it was eliminated from the feedback by filtering and noise masking. Obtained estimates varied between 3 and 5 ms, but were not systematically related to the experimental manipulation.


The speech intelligibility index (SII) is frequently used to predict the speech intelligibility for speech in a given interfering noise. However, the SII model only has been validated for speech in stationary noise. Since the SII degrades from speech and noise spectra, it does not take into account any fluctuations in the masking noise. Hence, the model will yield similar SII values, regardless of the degree of fluctuation. In contrast, from the literature it is clear that normal-hearing listeners can benefit from the fluctuations in the noise. The present paper describes an SII-based approach to model speech reception thresholds (SRTs) for speech in both stationary and fluctuating noise. The basic principle of this approach is that both speech and noise signals are partitioned into small time frames. Within each time frame, the conventional SII is determined, yielding the speech information available to the listener at that time frame. Next, the SII values of these time frames are averaged, resulting in the SII for that particular condition. With the aid of SRT data from the literature, it will be shown that this approach can give a good account for most existing data.

LaSC19. Direct measurement of single and multiple passband intelligibilities: Comparison with estimates based upon the Speech Intelligibility Index. Richard M. Warren, James A. Bashford, Jr., and Peter W. Lenz (Dept. of Psych., Univ. of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201, mwaffen@uw.edu)

The intelligibility of individual bands spanning the speech spectrum is of interest for theoretical and practical reasons, and has been the subject of considerable experimental investigation. Direct measurement of passband intelligibility can be confounded with contributions from filter slopes, but by employing sufficiently high orders of FIR filtering, the present study has removed all slope contributions and measured directly intelligibilities of 1 octave and 1/3-octave passbands. Stimuli employed were based upon the same commercial recording of monosyllabic words and the same frequency bands used for the Speech Intelligibility Index (SII) [American National Standards Institute, S3.5, 1997]. SII employs an indirect procedure for estimating intelligibility: lists of band “importance” values are derived from intelligibility scores for high-pass and low-pass speech having incrementally varied cutoff frequencies. These importance values are correlated with intelligibility, and were transformed into intelligibility estimates using the published transfer function. Directly measured intelligibilities differ for some, but not all, SII-based intelligibility estimates for bands heard singly and in combination. Direct determination of intelligibilities of individual and multiple passbands is suggested as a simple and accurate alternative to the methods based upon SII and other indirect procedures for estimating the intelligibility of frequency-limited speech. [Work supported by NIH.]
This paper examines the interaction of language-general, signal enhancement strategies and language-specific, phonological enhancement strategies in clear speech production and perception in Croatian, a language with a phonemic vowel length contrast and a relatively small inventory of five vowel qualities. Two native speakers of Croatian (one male, one female) read 20 nonsense sentences in conversational and clear speech. In a sentence-in-noise perception test, native Croatian listeners more accurately recognized words produced in clear than in conversational speech by both talkers. However, this clear speech intelligibility benefit was greater for the male than the female talker. Acoustic analyses showed that in clear speech both talkers enhanced the overall acoustic salience of the signal (slower, more frequent pauses, wider pitch range) and expanded the vowel space, although the male showed slightly less extensive vowel space expansion. However, only the male talker enhanced the phonemic vowel length contrast, suggesting that the observed asymmetry in the clear speech intelligibility benefit between the talkers may be due to the contribution of this phonological enhancement feature of his (but not her) clear speech production. The results suggest that speech intelligibility is most effectively enhanced by acoustic-phonetic modifications that reflect a combination of acoustic-auditory and phonological factors.

Vowels in clear and conversational speech: Talker differences in intelligibility for elderly hearing-impaired listeners. Sarah Hargus Ferguson (Dept. of Speech-Lang.-Hearing, Univ. of Kansas, Dole Ctr., 1000 Sunnyside Ave., Rm. 3001, Lawrence, KS 66045, safergus@ku.edu)

Several studies have shown that when a talker is instructed to speak as though talking to a hearing-impaired person, the resulting clear speech is significantly more intelligible than typical conversational speech. A recent study of 41 talkers [S. H. Ferguson and D. Kewley-Port, J. Acoust. Soc. Am. 111, 2482 (2002)] demonstrated that for normal-hearing listeners identifying vowels in noise, the amount of improvement talkers achieve by speaking clearly varies widely. Acoustic analyses of 12 of those talkers suggested that the amount of clear speech vowel intelligibility benefit was related to specific clear speech acoustic strategies adopted by the talkers. However, an earlier paper [S. H. Ferguson and D. Kewley-Port, J. Acoust. Soc. Am. 112, 256–271 (2002)] suggested that clear speech acoustic strategies that improve vowel intelligibility for normal-hearing listeners may not be beneficial for the group of listeners for whom clear speech is actually intended: listeners with hearing loss. The current project will explore this issue by assessing vowel intelligibility in clear and conversational speech for older hearing-impaired adults, using materials from the 12 talkers cited above. [Work supported by the University of Kansas Center for Research New Faculty General Research Fund and by NIHDCD-02229.]

Verbal transformation effect and the neighborhood activation model: Exploring the boundaries of the neighborhood. Peter W. Lenz, Richard M. Warren, and James A. Bashford, Jr. (Dept. of Psych., Univ. of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201, plenz@uwm.edu)

A recorded word repeating over and over undergoes a succession of illusory changes to other words and syllables in the listener’s lexicon, as well as to nonwords. This verbal transformation effect (VTE) appears to involve successive satiations of a dominant representation and serial replacements by competing representations. Early during the presentation of a VTE stimulus, reported illusory forms are typically lexical and nonlexical neighbors of the veridical stimulus (i.e., forms differing from the stimulus by a single phoneme). Interestingly, presentation of a lexical stimulus initially evokes far more reports of nonlexical neighbors while a nonlexical stimulus evokes predominantly lexical neighbors. After 6 to 7 min, the perceived forms are no longer immediate neighbors of the veridical stimulus, and differ by two phonemes on average. The neighborhood activation model (NAM) considers that identification of spoken words involves the activation of competing, phonetically similar lexical and nonlexical representations. Activation of one of these competitors reaches a critical level and that competitor enters awareness. The VTE polls the population of activated representations, providing a means of defining the functional boundaries of neighborhoods as well as the dynamics of competitor interaction. [Work supported by NIH.]

Frequency effects in phoneme processing. Danny Moates, Verna Stockmal, and Zinny Bond (Ohio Univ., Athens, OH 45701)

This study tested the hypothesis that phonological segments are activated during word recognition in proportion to their frequency of use, analogous to frequency effects in whole word recognition. Preliminary evidence for the hypothesis was given by Moates, Bond, and Stockmal [LabPhon 7 (2002)] using a word reconstruction task. The present study used a gating task in which progressively longer fragments of a word are presented to listeners who must identify the word after as few gates as possible. High- and low-frequency segments were contrasted by presenting them in word pairs that differed in two segments, e.g., collision–collusion, where /lu/ is more frequent than /lu/. We constructed 15 word pairs contrasting vowels and 16 pairs contrasting consonants (e.g., relief–release, where /sl/ is more frequent than /fl/). Identification judgments were gathered from 125 participants. An ANOVA showed high-frequency consonants to be identified at significantly earlier gates than their matched low-frequency consonants with both subjects and items as random factors. No such effect appeared for vowels. Also, whole words containing the high-frequency segments were not identified significantly earlier than those containing low-frequency segments. If the phoneme frequency effect is reliable, then spoken word recognition models should address it.

Adult perceptions of phonotactic violations in Japanese. Laurel Fais (Dept. of Psych., Univ. of British Columbia, 1401-2136 West Mall, Vancouver, BC V6T 1Z4, Canada, fais@psych.ubc.ca), Sachyio Kajitakawa (NTT Commun. Sci. Labs., Seika-cho, Kyoto 619-0237, Japan), Janet Werker (Univ. of British Columbia, Vancouver, BC, Canada), and Shigeaki Amano (NTT Commun. Sci. Lab.)

Adult Japanese speakers “hear” epenthetic vowels in productions of Japanese-like words that violate the canonical CVCCVCV form by containing internal consonant clusters (CVCCVC) [Dupoux et al., J. Exp. Psychol. 25, 1568–1578 (1999)]. Given this finding, this research examined how Japanese adults rated the goodness of Japanese-like words produced without a vowel in the final syllable (CVC), and words produced without vowels in the penultimate and final syllables (CVCC). Furthermore, in some of these contexts, voiceless vowels may appear in fluent, casual Japanese productions, especially in the Kanto dialect, and in some, such voiceless vowels may not appear. Results indicate that both Kanto and Kinki speakers rated CVC productions for contexts in which voiceless vowels are not allowed as the worst; they rated CVC and CVCC contexts in which voiceless vowel productions are allowed as better. In these latter contexts, the CVC words, which result from the loss of one, final, vowel, are judged to be better than the CVCC words, which result from the loss of two (final and penultimate) vowels. These results mirror the relative seriousness of the phonotactic violations and indicate listeners have tacit knowledge of these regularities in their language.

Can Japanese listeners recognize phonemes as well as morae as word boundary cues? Takashi Otake (Dokkyo Univ., 1-1, Gakuen-cho, Soka, Saitama 340-0042, Japan, otake@dokkyo.ac.jp) and Yoko Sakamoto (Grad. School of Foreign Lang., Dokkyo Univ., Japan)

According to Shortlist (Norris et al., 1995), one of the most recent universal models in spoken word recognition, metrical segmentation strategy, provides us with important cues on word boundaries, which are determined by linguistic rhythm. If this model is applied to Japanese, the cue...
must be morae because it is a mora-timed language. One thing which is not fully understood yet is whether or not a phoneme boundary exists in Japanese. According to Shrintlist, a possible word constraint (PWC) predicts that impossible words are inhibited, so that a phoneme boundary should not exist in Japanese. In order to test this hypothesis, two experiments were conducted using a modified version of word spotting task (search ase and hamu in nip#ase and ri#hamu) with three groups of Japanese subjects, 19 adults, 35 children, with or without Roman alphabet. The miss rates show that the subject groups could successfully find embedded words beyond both phoneme (adults:15%; children with alphabetic knowledge: 17%; children without it: 23%) and mora (adults:7%; children with alphabetic knowledge: 9%; children without it: 13%) boundaries with high accuracy, suggesting that in fact Japanese could be sensitive to phoneme boundaries, but the PWC inhibits them because they are not possible words.

1aSC26. Developmental changes in the perception of syllable-final stop voicing are not explained by age-related differences in auditory sensitivity. Susan Nittouer (Ctr. for Persons with Disabilities, Utah State Univ., 6840 Old Main Hill, Logan, UT 84322)

Two cues to voicing decisions for syllable-final stops are vocolic du-

ration and formant transitions, especially F1, at syllable offset. Develop-

mental studies show that children weight vocolic duration less and formant
transitions more than adults. This study tested the hypothesis that age-

related differences in auditory sensitivity can explain these developmental
changes by measuring adults’ and children’s sensitivities to changes in the
duration of nonspeech, complex tones and analogous speech stimuli. The
durations of three-tone sinusoidal stimuli and natural cob and cop were
varied. Duration difference limens (DLs) were similar for listeners across
ages, and similar for nonspeech and speech. Also, the extent of frequency
fall was varied in three-tone sinusoidal stimuli for the first tone (T1) only
or for all three tones, as well as for F1 only in synthetic buck/bug. DLs for
frequency extent were similar for listeners across ages for nonspeech
stimuli when all three tones fell, but children had larger DLs than adults
when only T1 or F1 fell, a finding contrary to labeling results showing
that children weight this very property more than adults. Overall the evi-
dence contradicted the hypothesis that age-related differences in auditory
sensitivity can explain developmental changes in the perception of
syllable-final stop voicing.

1aSC27. Implicit learning of nonadjacent phonotactic dependencies in the perception of spoken language. Conor T. McLennan and Paul A. Luce (Lang. Percept. Lab., Dept. of Psych. and Ctr. for Cognit. Sci., Univ. at Buffalo, 245 Park Hall, Buffalo, NY 14260, mclennan@buffalo.edu)

We investigated the learning of nonadjacent phonotactic dependencies in
adults. Following previous research examining learning of dependen-
cies at a grammatical level (Gomez, 2002), we manipulated the co-
ocurrence of nonadjacent phonological segments within a spoken sy-
lable. Each listener was exposed to consonant-vowel-consonant nonword
stimuli produced by one of two phonological grammars. Both languages
contained the same adjacent dependencies between the initial consonant-
vowel and final vowel-consonant sequences but differed on the co-
occurrents of initial and final consonants. The number of possible types
of vowels that intervened between the initial and final consonants was also
manipulated. Listeners learning of nonadjacent segmental dependencies
were evaluated in a speeded recognition task in which they heard (1) old
nonwords on which they had been trained, (2) new nonwords generated by
the grammar on which they had been trained, and (3) new nonwords
generated by the grammar on which they had not been trained. The results
provide evidence for listener’s sensitivity to nonadjacent dependencies.
However, this sensitivity is manifested as an inhibitory competition effect
rather than a facilitative effect on pattern processing. [Research supported by Research Grant No. R01 DC 0265802 from the National Institute on
Deafness and Other Communication Disorders, National Institutes of
Health.]

1aSC28. Production frequency effects in perception of phonological variation. Cynthia M. Connine and Larissa J. Ranbom (Dept. of Psych., Binghamton Univ., P.O. Box 6000, Binghamton, NY 13702, connine@binghamton.edu)

Two experiments were conducted that investigated the relationship be-
tween phonological variant occurrence frequency (based on a corpus
analysis of conversational speech) and auditory word recognition. The
variant investigation was an alternation between the presence of [nt] and a
nasal flap (e.g., center, cen’er). The corpus analysis showed that 80% of
productions are nasal flaps, with wide variability across words (from 6%
for “enter” to 100% for “twenty”). In a production goodness rating ex-
periment, listeners rated [nt] productions as better than their nasal flap
counterparts. For individual items, a strong positive correlation was found
between nasal flap frequency and goodness ratings: words typically pro-
duced with nasal flaps were rated as better productions. A lexical decision
experiment showed that nasal flap variants were recognized more slowly
and less accurately than [nt] versions. The rated quality of the nasal-
flapped production was strongly correlated with the results of the lexical
decision task: nasal-flapped words considered highly acceptable were rec-
ognized more quickly and accurately than words rated as poor nasal flap
productions. The results demonstrate a strong relationship between expe-
rienced variant frequency and auditory word recognition and suggest that
phonological variation is explicitly represented in the mental lexicon.

1aSC29. Effects of syllable onset length in determining word-likeness. Jordan Brewer, Benjamin V. Tucker, and Michael Hammond (Dept. of Linguist., Univ. of Arizona, P.O. Box 210028, 1100 E. University Blvd., Tucson, AZ 85721-0028)

Previous research on word-likeness has determined effects of neigh-
bors, and frequency of word parts (Bailey and Hahn, 2001; Coleman and
Pierrehumbert, 1997; Frisch et al., 2000). Holding those contributors con-
stant, we conducted an experiment designed to determine if there is also
an independent effect of syllable onset length in determining word-
likeness. Twenty students, native speakers of English, participated in a
lexical decision task in which they were asked to rank tokens, e.g., /vork/,
/strork/ with the syllable onset varied between one and three
segments) from 1 to 7. A 7 indicated a word that does not sound at all like
a possible English word and 1 indicated a word that sounds like it could be
a real English word (that you never heard before). A significant effect was
found of syllable onset length $F(3,20)=4.121; p<0.024$. Specifically,
a significant effect was found by subject but not by item between onsets of
one segment versus two $F(2,20)=6.362; p<0.021$, as well as one seg-
ment versus three $F(2,20)=7.436; p<0.013$, but there was no effect of
two versus three $F(2,20)=0.019$.

1aSC30. Auditory phonological priming in children and adults during word repetition. Miranda Cleary and Richard G. Schwartz (City Univ. of New York Grad. Ctr., Speech and Hearing Sci., 365 Fifth Ave., New York, NY 10016, mcleary@gc.cuny.edu)

Short-term auditory phonological priming effects involve changes in
the speed with which words are processed by a listener as a function of
recent exposure to other similar-sounding words. Activation of phonological/lexical representations appears to persist beyond the imme-
diate offset of a word, influencing subsequent processing. Priming effects
are commonly cited as demonstrating concurrent activation of word/
phonological candidates during word identification. Phonological priming
is controversial, the direction of effects (facilitating versus slowing) vary-
ing with the prime-target relationship. In adults, it has repeatedly been
demonstrated, however, that hearing a prime word that rhymes with the
following target word (BS1=50 ms) decreases the time necessary to initiate
repetition of the target, relative to when the prime and target have no
phonemic overlap. Activation of phonological representations in children
has not typically been studied using this paradigm, auditory-word +
picture-naming tasks being used instead. The present study employed an
auditory phonological priming paradigm being developed for use with
normal-hearing and hearing-impaired children. Initial results from normal-hearing adults replicate previous reports of faster naming times for targets following a rhyming prime word than for targets following a prime having no phonemes in common. Results from normal-hearing children will also be reported. [Work supported by NIH-NIDCD T32DC000399.]

LaSC31. F0 peaks aligned with nonprominent syllables in American English. Stefanie Shattuck-Hufnagel (RLE, MIT, 77 Massachusetts Ave., Cambridge, MA 02139; stef@speech.mit.edu), Nanette Veilleux (Simmons College), Alejna Brugos (Boston Univ.), and Robert Speer (RLE, MIT)

The occurrence of F0 peaks on nonprominent syllables in American English (e.g., -ing or a- in reading again) raises the question of how to label these inflection points. This pattern is not infrequent, as shown by samples from two prosodically labeled corpora of natural speech (ToBI labeled MIT Maptask and BU FM Radio News). The Maptask sample from a single speaker (235 seconds, 520 words) contained 46 H* !H* sequences; 19 had an F0 peak on a weak syllable between the two accented syllables; Individual speakers vary in their use of this pattern: for 3 Radio News speakers reading the same paragraph, each speaker produced 11 H* !H* sequences of which 6, 1, and 5, respectively, showed the non-accent-aligned peak. Informal listening as well as experiments described by Dilley (2004) suggest that alignment of the F0 peak with different nonprominent syllables between the two accents can change the perceived relative prominences of the accents, but current ToBI labels do not specify this alignment. Because F0 inflection points unaligned with the syllables perceived as prominent are common, and their location influences perceived prominence patterns, it is important to specify this aspect of an intonational contour. [Work supported by NIH Grants DC0075, DC02978, DC02125.]

LaSC32. The timing of speech-accompanying gestures with respect to prosody. Margaret Renwick, Stefanie Shattuck-Hufnagel, and Yelena Yasinik (Speech Commun. Group, 36-511 MIT, 77 Massachusetts Ave., Cambridge, MA)

The question of how and whether the body movements that accompany speaking are timed with respect to the speech has often been studied, and investigators have reached different conclusions depending on the types of gestures and aspects of prosody attended to. The ToBI system for labeling pitch accents (phrase-level prosodic prominences) and intonational phrase boundaries, which provides a well-defined inventory of prosodic elements, was used to label several sound files from videotaped lectures in English. A particular type of gesture, i.e., discrete sharp rapid movements that reach a perceptually salient end point, was separately labeled for syllable location in the visual display of the same lecture samples. Preliminary analysis showed a strong correlation between this type of "stroke-like" gesture of the head or hands and pitch accented syllables. For example, for one speaker of Australian English, 168 of 195 stroke-like gestures (86%) occurred with a pitch-accented syllable. If these observations from coarse-grained temporal labeling are confirmed by the frame-by-frame labeling now under way, it will suggest that the study of speech-accompanying gestures can provide evidence for the prosodic structure of spoken utterances, and raise the possibility that a complete model of speech production planning should include a gestural component.

LaSC33. Asking questions with focus. Fang Liu (Dept. of Linguist., Univ. of Chicago, 1010 E. 59th St., Chicago, IL 60637, liufang@uchicago.edu) and Yi Xu (Haskins Labs., New Haven, CT 06511)

This study investigates how different interrogative meanings interact with focus in determining the overall F0 profile of a question. We recorded eight native speakers of Mandarin producing statements, yes–no questions with and without a question particle, wh questions, incredulous questions, and confirmation questions. In each sentence, either the initial, medial, final, or no word was focused. The tonal components of the sentences are all high, all rising, all low, or all falling. F0 contours were extracted by measuring every complete vowel period in the initial, medial, and final disyllabic words in each sentence. Preliminary results show that in both statements and questions, the pitch range of the focused words is expanded and that of the postfocus words suppressed (compressed and lowered). However, postfocus pitch-range suppression seems less extensive in questions than in statements, and in some question types than in others. Finally, an extra F0 rise is often observed in the final syllable of a question unless the syllable is the question particle which has the neutral tone. This is indicative of a high or rising boundary tone associated with the interrogative meaning, which seems to be superimposed on the tone of the sentence-final syllable. [Work supported by NIDCD DC03902.]

LaSC34. Multiple effects of consonant manner of articulation and intonation type on F0 in English. Yi Xu and Andrew Wallace (Haskins Labs., 270 Crown St., New Haven, CT 06511, xu@haskins.yale.edu)

In this study we examine how consonant manner of articulation interacts with intonation type in shaping the F0 contours in English. Native speakers of American English read aloud words differing in vowel length, consonant manner of articulation and consonant position in word. They produced each word in either a statement or question carrier. F0 contours of their speech were extracted by measuring every complete vowel period. Preliminary results based on graphic analysis of three speakers data suggest that there are three distinct consonantal effects: F0 interruption due to devoicing, a large but brief (10–40 ms) F0 raising at the onset of voicing, and a smaller but longer-lasting F0 raising throughout a large proportion of the preceding and following vowels. These effects appear to be imposed on a continuously changing F0 curve that is either rising-falling or falling-rising, depending on whether the carrier sentence is a statement or a question. Further analysis will test the hypothesis that these continuous curves result from local pitch targets that are assigned to individual syllables and implemented with them in synchrony regardless of their segmental composition. [Work supported by NIDCD Grant No. R01 DC03902.]

LaSC35. Acoustic correlates of perceived rhythm in spoken English. J. Devin McAuley (Dept. of Psych., Bowling Green State Univ., Bowling Green, OH 43403) and Laura C. Dilley (Harvard Univ., Cambridge, MA 02139)

Two experiments examined the relationship between speech timing and perceived rhythm for a corpus of approximately 900 spoken sentences. In experiment 1, trained listeners applied an annotation system for perceptual isochrony to the corpus. For each sentence, listeners assigned beats to syllables and judged whether the intervals between successive beats were equal or unequal, permitting the identification of perceptually isochronous speech fragments (or beat chains). For each such beat chain, the intervals between vowel onsets of successive beat syllables were determined. Overall, there was good agreement among labelers about what constituted a beat chain, with the average inter-beat-interval equal to approximately 550 ms. In experiment 2, naive listeners rated the rhythmicity of each sentence on a 6-point scale, ranging from 1 (very nonrhythmic) to 6 (very rhythmic). Sentences were judged as more rhythmic when the longest identified beat chain in experiment 1 for that sentence contained more beats and had an average IBI closer to 550 ms. IBI variability was not a significant predictor of perceived rhythmicity. The results of both
experiments will be discussed in terms of preferred tempi in perceived rhythm and entrainment models of auditory event timing (McAuley and Jones, 2003).

1aSC36. The “listener” in the modeling of speech prosody. Klaus J. Kohler (Inst. of Phonet. and Digital Speech Processing (IPDS), Univ. of Kiel, D-24098 Kiel, Germany, kjk@ipds.uni-kiel.de)

Autosegmental-metrical modeling of speech prosody is principally speaker-oriented. The production of pitch patterns, in systematic lab speech experiments as well as in spontaneous speech corpora, is analyzed in F0 tracings, from which sequences of H(gh) and L(ow) are abstracted. The perceptual relevance of these pitch categories in the transmission from speakers to listeners is largely not conceptualized; thus their modeling in speech communication lacks an essential component. In the metalinguistic task of labeling speech data with the annotation system ToBI, the “listener” plays a subordinate role as well: H and L, being suggestive of signal values, are allocated with reference to F0 curves and little or no concern for perceptual classification by the trained labeler. The seriousness of this theoretical gap in the modeling of speech prosody is demonstrated by experimental data concerning F0-peak alignment. A number of papers in JASA have dealt with this topic from the point of synchronizing F0 with the vocal tract time course in acoustic output. However, perceptual experiments within the Kiel intonation model show that “early,” “medial” and “late” peak alignments need to be defined perceptually and that in doing so microprosodic variation has to be filtered out from the surface signal.

1aSC37. Prosodic complexity and phrase length as factors in pause duration. Jelena Krivokapic (Linguist. Dept., Univ. of Southern California, 3601 Watt Way, Grace Ford Salvatori Hall 301, Los Angeles, CA 90089-1693, krivokap@usc.edu)

Research on pauses has mainly focused on predicting the likelihood of pause occurrence and on the effect of syntactic structure on pause duration within an utterance. Very little is known about what factors, apart from syntactic and discourse factors, influence the length of pauses between utterances or phrases. This experiment examines the effect of prosodic structure and phrase length on pause duration. Subjects read 24 English sentences varying along the following parameters: (a) the length in syllables of the intonational phrase preceding and following the pause and (b) the prosodic structure of the intonational phrase preceding and following the pause, specifically whether or not the intonational phrase branches into smaller phrases. In order to minimize variability due to speech rate and individual differences, speakers read sentences synchronously in dyads (Cummins, 2002; Zvonik and Cummins, 2002). The results show that length has a significant effect on pause duration both pre- and postboundary for all dyads, and that prosodic complexity has a significant postboundary effect for some dyads. The possible reasons for the observed pause duration effects and the implications of these results on the question of incrementality in speech production are discussed. [Work supported by NIH DC03172.]

1aSC38. Distinct relative F0 levels elicit categorical effects in F0 maximum and minimum alignment. Laura C. Dilley and Meredith Brown (MIT Speech Commun. Group, 50 Vassar St., 36-549, Cambridge, MA 02139, dilley@mit.edu)

A standard assumption in intonation research is that the presence and timing of fundamental frequency (F0) maxima and minima relative to segments are crucial phonetic characteristics of intonation patterns. The present experiment tests an alternative hypothesis that the representation of intonation patterns is based on the relative pitch levels of syllables in sequence. Synthetic stimuli were created using the phrase Some lemonade with an overall rising/falling or falling/rising intonation pattern. Cues to the presence and timing of F0 maxima and minima were eliminated by replacing the F0 across lemon- with level F0 contours and replacing the sonorant consonants before and after each target vowel nucleus with Gaussian noise. Four continua were created by shifting the F0 levels of one or both syllables in equal 0.5- or 0.75-semitone increments. Thirteen subjects imitated randomized stimuli presented over headphones. Results showed that alignment of maxima and minima in imitations was predictably related to relative F0 level: F0 maxima (minima) were aligned early in the segmental string when the first target syllable in the stimulus was higher (lower) than the following syllable; otherwise, maxima and minima were aligned late. These results provide support for models of intonation based on relative pitch levels.

1aSC39. Prosodic effects on glide-vowel sequences in three Romance languages. Ioana Chitoran (Dartmouth College Linguist., HB 6087, Hanover, NH 03755)

Glide-vowel sequences occur in many Romance languages. In some they can vary in production, ranging from diphthongal pronunciation [ia-ie] to hiatus [i-a]. According to native speakers’ impressionistic perceptions, Spanish and Romanian both exhibit this variation, but to different degrees. Spanish favors glide-vowel sequences, while Romanian favors hiatus, occasionally resulting in different pronunciations of the same items: Spanish (b[i]e-la, [i-d]ana), Romanian (b[i]-e-la, [i-d]ana). The third language, French, has glide-vowel sequences consistently (b[j]elle). This study tested the effect of position in the word on the acoustic duration of the sequences. Shorter duration indicates diphthong production [jV], while longer duration, hiatus [IV]. Eleven speakers (4 Spanish, 4 Romanian, 3 French), were recorded. Spanish and Romanian showed a word position effect. Word-initial sequences were significantly longer than word-medial ones (p<0.001), consistent with native speakers more frequent description of hiatus word-initially than medially. The effect was not found in French (p>0.05). In the Spanish and Romanian sentences, V in the sequence bears pitch accent, but not in French. It is therefore possible that duration is sensitive not to the presence/absence of the word boundary, but to its position relative to pitch accent. The results suggest that the word position effect is crucially enhanced by pitch accent on V.


Seven British English dialects were studied to see what prosodic distinctions are made between statements and questions in read speech. A set of Bayesian classifiers was built upon feature vectors obtained from a spectral analysis of measures of (1) F0, (2) loudness, (3) spectral slope and (4) voicing periodicity. It was found that the prosodic information useful for the question/statement distinction is distributed broadly across the utterance, and that loudness and spectral slope can be nearly as informative as F0 (voicing is less informative). The three important acoustic features carry somewhat less than one bit of information each, so prosodic information could be valuable to the listener, and the listener may be able to make the question/statement decision even early in the utterance. The contrast differs from one acoustic property to the next: F0 is marked primarily by slow variations. Conversely, the spectral slope and loudness measurements primarily use shorter-wavelength features, corresponding to structures that are syllable or two long. We also find substantial differences in the prosodic information that different dialects use, and substantial differences between speakers of the same dialect. [Research supported by the UK Economic and Social Research Council, Grant RES 00-23-1049.]
Underwater Acoustics and Acoustical Oceanography: Sediment Acoustics and Scattering

Nicholas P. Chotiros, Chair

Office of Naval Research, Code 3210A, 800 North Quincy Street, Arlington, Virginia 22217-5660

Chair’s Introduction—7:40

Contributed Papers

7:45

1aUW1. An experimental verification of the spherical wave effect in unconsolidated sediments. H. John Camin and Marcia Isakson (Appl. Res. Labs., Univ. of Texas, 10000 Burnet Rd., Austin, TX 78758, misakson@arlut.utexas.edu)

Several research institutes are investigating models to properly describe high-frequency acoustic properties of unconsolidated sediments. This talk will focus on comparing experimental reflection coefficient measurements over a range of frequency (30–160 kHz) and angle to five of these theoretical sediment models: the fluid/liquid model, the visco-elastic model, the effective density fluid model (EDFM), Buckingham’s microsliding model, and the Biot/Stoll poro-elastic model. Since spherical transducers were used in the experiment, spherical wave effects will be added to the models using a plane-wave decomposition method. A common set of parameters will be used in all models. Fluid and visco-elastic model parameters have been measured and verified at the test location. The Biot parameters were determined from measured values and previous inversions of plane-wave reflection data taken at the same location. Air/water interface data will be measured and computed for each model as a test case. [Work supported by ONR, Ocean Acoustics.]

8:00

1aUW2. Low-frequency sound speed reduction in unconsolidated sediments by grain rolling and sliding. Marcia Isakson and Nicholas Chotiros (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029)

In recent experiments in the Gulf of Mexico, measured low-frequency sound speeds in sandy ocean sediments were significantly less than those predicted by Woods equation, the theoretical lower limit for sound speed in unconsolidated sediment [Williams et al., IEEE J. Ocean Eng. 27(3), 413–428; Turgut and Yamamoto, J. Acoust. Soc. Am. 87, 2376–2382 (1990)]. Therefore, the conventional treatment of sound wave velocity is apparently overlooking processes necessary to describe wave propagation in unconsolidated sediments. One possible process is a partitioning of sound wave energy into grain rolling and sliding, which leads to additional terms in the kinetic energy equation for the sand grain frame. This additional kinetic energy can be treated as an incremental virtual mass. The virtual mass terms can be evaluated given the scintillation index of the sound speed in unconsolidated sediments by grain rolling and sliding. This work will present the derivation of the virtual mass terms, calculations based on likely parameter values and a comparison with the experimental measurements. [Work supported by ONR, Ocean Acoustics.]

8:15

1aUW3. Field experiments on the velocity dispersion and attenuation of acoustic waves through sandy sediments in shallow waters. Tokuo Yamamoto, Haruhiko Yamaoka (AMP, RSMAS, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33176), and Junichi Sakakibara (JFE Civil Corp., Taito-ku, Tokyo, Japan)

The velocity dispersion and attenuation of acoustic waves though a sandy bottom in a shallow ocean were measured for a broad frequency band of 1–500 kHz using a buried horizontal hydrophone array and several broadband piezoelectric sources. The sands have a fairly uniform grain size of 0.22 mm and a porosity of 0.44. The experimental data were compared with the Biot theory, the BISQ theory, the squirt flow theory and the patchy-saturated theory. The data—theory comparisons indicate that the Biot theory best predicts the acoustic propagation though the sand bottom in shallow water in the medium to high frequency band of 1–500 kHz. The data—theory comparisons also confirm that the permeability, porosity, and added mass coefficient of shallow water sediments can be extracted from acoustic data. [Work supported by ONR Code 321 OA.]

8:30

1aUW4. Inversion of sediment parameters in Biot parameter space. Gopi R. Potti and James H. Miller (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882)

Long range sediment tomography technique [Potti et al., J. Acoust. Soc. Am. 108, 973–986 (2000)] using explosive sources has been developed and applied to data from the Primer and ASIAEX experiments. This inversion scheme estimates the compressional wave speed and attenuation matching the individual modal arrival times, i.e., inversion is carried out in the geoaoustic parameter space. This inversion scheme is now modified to invert for parameters in the sediment physical property space. Using the Biot–Stoll model as the interface between sediment physical and geoaoustic spaces, few of the most sensitive Biot parameters are inverted for. A sensitivity analysis will be carried out to identify the most sensitive Biot parameters and they will be considered as unknowns. Broad band acoustic data from wide band sources (WBS) collected on a vertical line array in frequency range 10–400 Hz will be used for the inversion. The goal of this study is to identify the sediment type (sand, silt, sandy-silt, clay, etc.) along with some physical attributes qualifying its nature (soft, hard, etc.) and the thickness of layers. This will enable us to describe the sediment characteristics over a large area with minimum number of parameters. [Work supported by ONR.]

8:45

1aUW5. Near-real-time geoacoustic characterization during the BOUNDARY 2003 experiment. Kevin D. Heaney (Lockheed-Martin ORINCON, 4350 N. Fairfax Dr., Ste. 470, Arlington, VA 22203, kevin.heaney@lmco.com) and Peter N. Nielsen (SACLANT Undersea Res. Ctr., La Spezia, Italy)

During the SAaCLANT Center BOUNDARY 2003 experiment on the Malta plateau, the rapid geoacoustic characterization (RGC) algorithm was applied. This algorithm uses the slope and spacing of shallow-water striations as well as the slope of the received level with range from a passing surface ship to estimate an effective geoacoustic sediment. By searching over a small set of representative sediments, using empirical relationships between grain size and compressional speed, density, and attenuation, the inversion can be carried out in near-real time. Ambient noise measurements of passing surface ships in the Malta plateau lead to geoacoustic inversion results obtained at sea that are comparable to previous work in the area where the inversions were obtained using vertical line arrays and controlled towed sources.
9:00
1aUW6. Environmental influences on the frequency dependence of effective bottom attenuation. James D. Nickila (Adv. Sonar Technol. Div. Naval Undersea Warfare Ctr., Newport, RI 02841), Kevin B. Smith (Naval Postgrad. School, Monterey, CA 93943, kbsmith@nps.navy.mil), and Gopu Potty (Univ. of Rhode Island, Narragansett, RI 02882)

Over the past several years, concern has grown over the appropriateness of bottom-attenuation models that assume a linear frequency dependence. Empirical analyses of experimental data have suggested power-law dependence with frequency exponents as high as 1.7 and above, but with large variability between geographic regions [Zhou et al., J. Acoust. Soc. Am. 82, 287–292 (1987)]. The fundamental cause of this dependence is unknown. In this analysis, the influence of the propagation and interaction with environmental variability is investigated. Specifically, a propagation model that assumes linear frequency dependence is employed which incorporates such environmental variability as range-dependent water column sound-speed profiles, bottom sound-speed gradients, bottom sound-speed and density fluctuations, and rough water/bottom interfaces. These data are then decomposed into normal-mode amplitudes at various ranges. Based on an approach similar to Pott et al. [J. Acoust. Soc. Am. 114, 1874–1887 (2003)], estimates of average bottom attenuation may be inverted directly. By computing such estimates over a band of frequencies, the effective frequency dependence of the bottom attenuation can be determined as a function of range. The dominant environmental influences will also be identified and quantified.

9:15
1aUW7. The frequency dependence of the relative sensitivities of the Biot parameters. Marcia Isakson and Traci Anne Neilson (Appl. Res. Labs., Univ. of Texas, 10000 Burnet Rd., Austin, TX 78758)

The frequency and angular dependence of the seabed reflection coefficients contain information about how the properties of the ocean bottom influence acoustic propagation. One description of the interaction between sound and sediment is the Biot–Stoll model, in which the sediment is parameterized by 13 properties. Some of these properties have a large influence on the reflection coefficients over a specific frequency band and little or no influence on reflection coefficients at other frequencies. The sensitivity of the reflection coefficients as a function of frequency to changes in the Biot–Stoll parameters are explored using a rotated coordinates inversion method. Rotated coordinates are eigenvectors obtained from an eigenvalue decomposition of the covariance matrix of gradients of the cost function to be optimized. In this case, the cost function is designed to minimize the difference between measured and modeled reflection coefficients. The resulting rotated coordinates and corresponding eigenvalues indicate the relative sensitivity of the reflection coefficients to changes in the parameters. The rotated coordinates inversion technique is applied to reflection coefficient data synthesized for a sandy bottom over subsets frequencies from 100 Hz to 1 GHz. [Work supported by ONR.]

9:30

Determination of the frequency behavior of modal attenuation coefficients is essential for estimation of propagation influences of poro-elastic sediments. A recent simplification of the Biot model [Pierce et al., J. Acoust. Soc. Am. 114, 2345 (2005)] provides an approach for this determination. For plane compressional waves in a homogeneous medium, the simplified theory reproduces the frequency-squared behavior of the full Biot model. However, data from a variety of shallow water locations suggest power-law exponents that are less than two. Inhomogeneities, particularly in the upper sediment layer, influence the frequency behavior. For a normally consolidated sediment (one never subjected to pressures higher than the current overburden pressure), depth profiles of porosity and sound speed can be estimated [Cederberg et al., J. Acoust. Soc. Am. 97, 2754–2766 (1995)]. We use these profiles to investigate the frequency behavior of the modal attenuations. Comparisons are provided with observations as well as previous numerical studies. [Work supported by ONR.]

9:45
1aUW9. Structure of acoustic reflection coefficient at the fluid-poroelastic interface and the medium propagating the slow wave. Keichi Ohkawa (Div. of Appl. Marine Phys., RSMAS, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149)

The closed-form expressions of the acoustic reflection coefficient at the fluid-poroelastic interface exist. Due to the complexity, it is, however, rather difficult to acquire good physical insight into dependencies of the coefficient on many measurable quantities defining the fluid/porous-medium interface. Introducing effective densities allows us to obtain fine physical insight into the structure of the reflection coefficient. In the absence of the shear wave, the fundamental structure of the reflection coefficient is of the form \( R = (1/Z_0 - 1/Z_1 - 1/Z_2)/(1/Z_0 + 1/Z_1 + 1/Z_2) \) where \( Z_0, Z_1, \) and \( Z_2 \) are impedances of sound, fast, and slow waves, respectively. At the normal incidence this gives the exact solution. The phase velocities of fast and slow waves, \( c_1 \) and \( c_2 \), are approximated as \( c_1 \approx \sqrt{K_1/\rho_1} \), \( c_2 \approx \sqrt{K_1/\rho_2} \), where \( C \) is the Biot’s coupling constant, \( K_1 \) is the fluid bulk modulus, \( \rho \) is the porosity, and \( \rho_1 \) and \( \rho_2 \) are effective densities of fast and slow waves. In air-saturated medium, \( \rho_2 \) approaches the dynamic added mass density, For the super fluid \( ^4 \)He, \( c_2 \) becomes the same result derived by Dutta. This supports that the medium propagating the slow wave is without doubt the pore fluid in any porous media.

10:00–10:15 Break

10:15
1aUW10. Approximate single wave equation derived from Biot’s porous media equations. Allan D. Pierce, William M. Carey (Boston Univ., Boston, MA 02215, adp@bu.edu), James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Mario Zampolli (NATO Undersea Res. Ctr., La Spezia, Italy)

Manipulation of Biot’s equations using low-frequency and long-wavelength approximations yields wave equation with a dissipation term that predicts plane wave attenuation coefficients proportional to the square of the frequency. Examination of Biot’s original derivation verifies that the wave speed in this equation is what results from Wood’s theory. The derivation predicts that the fluid displacement is a predictable fraction of that of the solid matrix, and this is consistent with the behavior in Brillouin’s model of an elastic spring loaded with masses of alternating magnitude. Biot’s assumption that the pores must be connected is argued to be unnecessary. Unconnected pores may be elongated with constrictions, and the sound field will cause some portions of the porelets to be compressed more than other portions, so that pressure differences across constrictions force viscous flow. A heuristic derivation shows that the wave attenuation coefficient is inversely proportional to viscosity and sound speed, and directly proportional to density, the number of porelets per unit volume, the average length of porelets, and the square of the cross-sectional area of the constrictions and frequency. The heuristic theory is supported with a detailed analysis of the energy dissipation within a dumbbell-shaped porelet within an extended isotropic elastic solid.

10:30
1aUW11. Resolution cell size effects on the statistics of seafloor backscatter. Shawn F. Johnson, Anthony P. Lyons, Doug Abraham (Appl. Res. Lab., The Pennsylvania State Univ., P.O. Box 30, State College, PA 16804), and Eric Pouliquen (NATO Undersea Res. Ctr., 19138 La Spezia, Italy)

An understanding of high-frequency scattering is essential for the design and implementation of successful object detection algorithms. Due to their highly heterogeneous nature, seafloors in shallow water pose a particularly difficult problem for target detection with scattered envelope
probability density function (PDF) often having heavier tails than the Rayleigh distribution. One aspect of high-frequency scatter that has not been resolved to date, and that should also shed light on the causes of non-Rayleigh distributions, is the relation between the size of the resolution cell of the sonar system and the envelope distributions. Seafloor scattering data from transmitted CW pulses (50 and 100 kHz) and FM signals (37–65 kHz) were collected by the NATO Undersea Research Centre off of Elba Island, Italy in May 2003. Water depths were less than 100 m and bottom structures of the test sites included Posidonia Oceanica sea grass, rippled sand, and rocks. Varying the length of the CW pulses and the bandwidth of the FM signals allowed the effective resolution cell size of the sonar to be varied. The impact of changing resolution cell size on the seafloor scattering statistics will be discussed as well as implications. [Research performed under ONR Grant No. N00014-03-1-0245.]

10:45


In a number of applications, including echo sounding of the ocean bottom, nondestructive evaluation of materials, and satellite altimetry, measurements of travel time and/or shape of a backscattered pulse are employed to retrieve information about position and properties of scatterers. When the scatterer is an extensive rough surface with sufficiently large slopes, the backscattered wave consists of numerous pulses, of which the earliest arrivals are often of primary interest. In this paper, the geometrical acoustics approximation is applied to study signal intensities and the difference between the travel times of waves scattered by a random rough surface and those reflected from the mean, plane boundary. For a Gaussian statistical distribution of roughness, probability density functions (PDFs) of the travel-time differences are found for the first and second arrivals in terms of variances of the surface roughness elevation and slope. The joint PDFs of the travel times and intensities are derived. It allows one to obtain the travel-time PDF for signals exceeding a given intensity threshold. It is shown that the travel time and the intensity are strongly correlated; on average, the earlier a signal arrives, the smaller is its intensity.

11:00


In SAX99 (Sediment Acoustics Experiment 1999), a number of high quality acoustic and environmental measurements were made in support of an ONR Department Research Initiative on high-frequency sound interaction in ocean sediments [M. D. Richardson et al. and E. I. Thoros et al., IEEE J. Ocean Eng. 26, 4–53 (2001)]. Supporting environmental measurements include digital stereo photography of the interface roughness which were converted [A. P. Lyons et al., IEEE J. Ocean Eng. 27, 515–524 (2002)] to digital elevation maps. Using these micro-bathymetry data sets, the measured backscatter strengths from SAX99 are modeled via an extension of the wedge assemblage model [R. S. Keiffer and J. C. Novarini, J. Acoust. Soc. Am. 107, 27–39 (2000)] that allows for density contrast between the water and sediment but assumes no change in the sound speed. In particular, SAX99 data in which changes in the acoustic backscatter can be directly related to the time evolution of the fine-scale interface morphology will be modeled and compared. [Work supported by ONR, Program Element No. 61153N-32 and by the DoD High Performance Computing Shared Resource Center (Stennis). This document has been reviewed and approved for public release.]

11:15

1aUW14. Scattering from marine sediments with discrete inclusions. Anatoly N. Ivakin (App. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, ivakin@apl.washington.edu)

A model of scattering in sediments with discrete inclusions is developed allowing prediction of bottom reverberation, given material parameters of the sediment and inclusions and their size distribution, which can be depth dependent. The frequency-angular dependencies of the seabed scattering strength for various types of the sediment and inclusions are calculated and discussed. An environmental data set obtained at SAX99 site (near Walton Beach, FL), including the sediment particle size distribution and roughness spectra, was used to compare contributions of volume and roughness components of the seafloor scattering. It is shown, in particular, that contribution of gravel and shell inclusions and coarse sand fraction in total scattering can be dominating (over roughness) at high frequencies (about 100 kHz and higher) and grazing angles above critical (about 30°) while roughness at SAX99 site is likely a dominating mechanism of bottom scattering at lower frequencies and grazing angles below critical. A combined model, taking into account both roughness and volume discrete scattering, is shown to be a good descriptor of bottom reverberation in a wide frequency-angular range. Possibilities for inversion of various sediment parameters from backscattering data are discussed. [Work supported by ONR, Ocean Acoustics.]

11:30

1aUW15. A rapid method for calculating time domain scattering from a penetrable bottom. C. Feuillade (Naval Res. Lab., Stennis Space Center, MS 39529-5004)

The density contrast surface is an appropriate model for many water-seafloor interfaces, and the formalism of D. Chu [J. Acoust. Soc. Am. 86, 1883–1896 (1989)], for the impulse response of a point source to a density contrast wedge, facilitates the extension of wedge assemblage boundary scattering models to rough penetrable surfaces. Chu’s method, however, is computationally intensive, since it necessitates finding the multiple roots of a characteristic eigenvalue equation, and evaluating a corresponding power series, for each wedge apex. In this work, the direct, reflected, and diffracted field components of a density contrast wedge are considered, and the relationships between them investigated. In particular, the origin and behavior of diffractions associated with specular reflections of the source in the faces of the wedge is studied. Results show how a proper physical understanding of these phenomena leads to a simple modification of the Biot–Tolstoy theory [J. Acoust. Soc. Am. 29, 381–391 (1957)], which then permits rapid calculations of acoustic bottom scattering and penetration in the time domain to be performed for typical sediment structures seen on the seafloor. [Work supported by ONR/Naval Research Laboratory]
sample at each station. Subsampling tubes were used to obtain undisturbed samples; laboratory measurements of density, compressional wave speed and attenuation, resistivity, and grain size were completed. For a small subset of samples, selected to represent a range of sediment types, measurements of permeability, shear wave speed, and attenuation were completed. [Research supported by ONR Grant No. N00014-00-1-0821.]

MONDAY AFTERNOON, 24 MAY 2004 ROYAL BALLROOM A, 2:30 TO 4:35 P.M.

Session 1pAAa

Architectural Acoustics: Multisensory Integration and the Concert Experience: How Visual Input Affects What We Hear

Jerald R. Hyde, Chair

Consultant on Acoustics, Box 55, St. Helena, California 94574

Chair’s Introduction—2:30

Invited Papers

2:35

1pAAa1. Multisensory integration and the concert experience: An overview of how visual stimuli can affect what we hear. Jerald R. Hyde (P.O. Box 55, St. Helena, CA 94574)

It is clear to those who “listen” to concert halls and evaluate their degree of acoustical success that it is quite difficult to separate the acoustical response at a given seat from the multi-modal perception of the whole event. Objective concert hall data have been collected for the purpose of finding a link with their related subjective evaluation and ultimately with the architectural correlates which produce the sound field. This exercise, while important, tends to miss the point that a concert or opera event utilizes all the senses of which the sound field and visual stimuli are both major contributors to the experience. Objective acoustical factors point to visual input as being significant in the perception of “acoustical intimacy” and with the perception of loudness versus distance in large halls. This paper will review the evidence of visual input as a factor in what we “hear” and introduce concepts of perceptual constancy, distance perception, static and dynamic visual stimuli, and the general process of the psychology of the integrated experience. A survey of acousticians on their opinions about the auditory-visual aspects of the concert hall experience will be presented. [Work supported in part from the Veneklasen Research Foundation and Veneklasen Associates.]

2:55

1pAAa2. Spectro-temporal interactions in auditory-visual perception: How the eyes modulate what the ears hear. Ken W. Grant (Walter Reed Army Medical Ctr., Army Audiol. and Speech Ctr., Washington, DC 20307-5001) and Virginie Van Wassenhove (Univ. of Maryland, College Park, MD 20742)

Auditory-visual speech perception has been shown repeatedly to be both more accurate and more robust than auditory speech perception. Attempts to explain these phenomena usually treat acoustic and visual speech information (i.e., accessed via speechreading) as though they were derived from independent processes. Recent electrophysiological (EEG) studies, however, suggest that visual speech processes may play a fundamental role in modulating the way we hear. For example, both the timing and amplitude of auditory-specific event-related potentials as recorded by EEG are systematically altered when speech stimuli are presented audiovisually as opposed to auditorily. In addition, the detection of a speech signal in noise is more readily accomplished when accompanied by video images of the speaker’s production, suggesting that the influence of vision on audition occurs quite early in the perception process. But the impact of visual cues on what we ultimately hear is not limited to speech. Our perceptions of loudness, timbre, and sound source location can also be influenced by visual cues. Thus, for speech and nonspeech stimuli alike, predicting a listener’s response to sound based on acoustic engineering principles alone may be misleading. Examples of acoustic-visual interactions will be presented which highlight the multisensory nature of our hearing experience.

3:15

1pAAa3. Basic research on auditory-visual interaction and listening in rooms. Frederic L. Wightman and Pavel A. Zahorik (Dept. of Psych. and Brain Sci. and Heuser Hearing Inst., Univ. of Louisville, Louisville, KY 40292)

Attending a live concert is a multisensory experience. In some cases it could be argued that hearing is the primary sense involved, but it is never the only one. Vision, smell, and even touch make important contributions to the overall experience. Moreover, the senses interact such that what one hears, for example, is influenced by what one sees, and vice-versa. This talk will address primarily the auditory aspects of the concert experience, focusing on the results of basic studies of human spatial hearing in reverberant environments, and how these results may help us understand the concert experience. The topics will include sound localization in anechoic and reverberant environments, the precedence effect, the cocktail party effect, the perception of distance, and the impact of...
The thinking which characterizes acoustics as a branch of physics and engineering has difficulty with the architectural design process—the process that generates a room concept in the imagination and experience of the architect. The architect has learned to “sense” the visual properties of a room as the design develops in the interaction between mind and media. Phrases such as “wanting to be” express the architectural intention but too often such intentions are dismissed as arbitrary; acoustics may then be about fixing the design with acoustical add-ons. Occasionally there is a true meeting of minds—a creative and receptive architect and an acoustician able to communicate at the level of the architectural intention. There is evidently an auditory dimension of wanting to be which is one with the visual. This paper explores the idea in several examples and concludes with suggestions for the training of acousticians.
Session 1pAAb

Architectural Acoustics and National Council of Acoustical Consultants: Student Design Competition (Poster Session)

Robert C. Coffeen, Cochair
School of Architecture and Urban Design, University of Kansas, Marvin Hall, Lawrence, Kansas 66045

Lily M. Wang, Cochair
Architectural Engineering, University of Nebraska–Lincoln, 200B Peter Kiewit Institute, 1110 South 67th Street, Omaha, Nebraska 68182-0681

Robin Glosemeyer, Chair
Jaffe Holden Acoustics, 501 Santa Monica Boulevard, Suite 606, Santa Monica, California 90401

The Technical Committee on Architectural Acoustics of the Acoustical Society of America and the National Council of Acoustical Consultants are sponsoring this Student Design Competition that will be professionally judged at this meeting.

The purpose of this design competition is to encourage students enrolled in architecture, architectural engineering, and other university curriculums that involve building design and/or acoustics to express their knowledge of architectural acoustics and building noise control in the schematic design of a building where acoustical considerations are of primary importance. This competition is open to undergraduate and graduate students from all nations.

The submitted designs will be displayed in this session and they will be judged by a panel of professional architects and acoustical consultants. Up to five entries will be selected for awards, one “First Honors” award and four “Commendation” awards. An award of $1,000 will be given to the entry judged “First Honors.” An award of $500 will be given to each of the entries judged “Commendation.”

Session 1pAB

Animal Bioacoustics: Natural Acoustic Behavior of Animals: Session in Memory of Donald R. Griffin II

Roderick A. Suthers, Chair
Medical Sciences, Indiana University, Bloomington, Indiana 47405

Invited Papers

2:15

1pAB1. Watching bats find food: Do we classify the signals, the strategies, or the bats?  James A. Simmons (Dept. of Neurosci., Brown Univ., Providence, RI 02912, james_simmons@brown.edu)

The fact that different families, genera, and even species of echolocating bats broadcast characteristic sonar signals has motivated numerous efforts to classify bats according to signal design, which has received support from correlations with both peripheral and central auditory physiology. Signal types vary according to the situations in which bats have been observed hunting for food, so this classification has been extended to the hunting strategies they use. The availability of new technical means for watching and documenting the behavior of echolocating bats in real time (thermal infrared video cameras, night-vision video with infrared illumination, video recorders with ultrasonic audio channels) makes it possible to follow individual bats long enough to observe variations in their behavior over periods of seconds to minutes. These observations reveal that at least some species nominally classified as using just one hunting strategy in fact use several strategies according to prevailing conditions, sometimes using different strategies in the course of only a few minutes. The historic inaccessibility of bats to real-time observation in the dark may have lead to exaggerated stereotyping of their behavior. [Work supported by ONR, NSF.]

2:35

1pAB2. Interaction of vestibular, echolocation, and visual modalities guiding flight by the big brown bat, Eptesicus fuscus.  Seth S. Horowitz (Dept. of Psychiatry, State Univ. of New York at Stony Brook, HSC T-10, Rm. 086, Stony Brook, NY 11794, shorowitz@neuropop.com) and James A. Simmons (Brown Univ., Providence, RI 02912)

The big brown bat (Eptesicus fuscus) is an aerial-feeding insectivorous species that relies on echolocation to avoid obstacles and to detect flying insects. Spatial perception in the dark using echolocation challenges the vestibular system to function without substantial visual input for orientation. IR thermal video recordings show the complexity of bat flights in the field and suggest a highly
dynamic role for the vestibular system in orientation and flight control. Laboratory studies of flight behavior under illuminated and dark conditions in both static and rotating obstacle tests were carried out while administering heavy water (D2O) to bats to impair their vestibular inputs. Eptesicus carried out complex maneuvers through both fixed arrays of wires and a rotating obstacle array using both vision and echolocation, or when guided by echolocation alone. When treated with D2O in combination with lack of visual cues, bats showed considerable decrements in performance. These data indicate that big brown bats use both vision and echolocation to provide spatial registration for head position information generated by the vestibular system.

2:55

1pAB3. Bat’s auditory system: Corticofugal feedback and plasticity. Nobuo Suga (Dept. of Biol., Washington Univ., One Brookings Dr., St. Louis, MO 63130)

The auditory system of the mustached bat consists of physiologically distinct subdivisions for processing different types of biosonar information. It was found that the corticofugal (descending) auditory system plays an important role in improving and adjusting auditory signal processing. Repetitive acoustic stimulation, cortical electrical stimulation or auditory fear conditioning evokes plastic changes of the central auditory system. The changes are based upon egocentric selection evoked by focused positive feedback associated with lateral inhibition. Focal electric stimulation of the auditory cortex evokes short-term changes in the auditory cortex and subcortical auditory nuclei. An increase in a cortical acetylcholine level during the electric stimulation changes the cortical changes from short-term to long-term. There are two types of plastic changes (reorganizations): centripetal best frequency shifts for expanded reorganization of a neural frequency map and centrifugal best frequency shifts for compressed reorganization of the map. Which changes occur depends on the balance between inhibition and facilitation. Expanded reorganization has been found in different sensory systems and different species of mammals, whereas compressed reorganization has been thus far found only in the auditory subsystems highly specialized for echolocation. The two types of reorganizations occur in both the frequency and time domains. [Work supported by NIDCO DC00175.]

3:15

1pAB4. Responses of inferior colliculus neurons are shaped by inhibitory projections from lower nuclei. George Pollak and Ruili Xie (Section of Neurobiology, Univ. of Texas, Austin, TX 78712, gpollak@mail.utexas.edu)

Bats are highly social animals that utilize a remarkably rich repertoire of signals for a variety of social interactions. This talk will explore how these signals are processed and represented in the inferior colliculus (IC) of Mexican free-tailed bats. Most IC neurons are selective for communication calls in that they respond only to some calls but not to others. The selectivity is due largely to inhibition. To evaluate the impact of the inhibitory projections from the dorsal or intermediate nucleus of the lateral lemniscus (DNLL and INLL), each nucleus was reversibly inactivated while responses of IC neurons to a suite of communication calls were recorded. Inactivation of both nuclei allowed these IC cells to respond to a larger number of calls, and thus become less selective than when the inhibitory innervation was intact. We therefore conclude that the inhibitory DNLL and INLL inputs to IC are functionally shaping the signal processing in the IC by suppressing some excitatory inputs. One consequence of the inhibition from these nuclei is that it shapes the selectivity of IC neurons for complex signals, allowing those IC cells to extract certain information from only some complex signals but not others. [Work supported by NIH Grant DC00268.]

3:35

1pAB5. Neural mechanisms underlying the analysis of sonar and social vocalizations: Spectral and temporal integration in the mustached bat. Jeffrey J. Wenstrup, Kiran Nataraj, Don Gans, and Kianoush Sheykholeslami (Dept. of Neurobiology, Northeastern Ohio Universities College of Medicine, 4209 State Rte. 44, Rootstown, OH 44272, jjw@neucom.edu)

The analysis of sonar echoes by the mustached bat depends on combination-sensitive neurons that respond best when distinct spectral elements of a sonar pulse and echo occur in a particular temporal relationship. Such integrative response properties underlie direct comparisons of acoustic features in outgoing pulses and returning echoes, comparisons thought to encode pulse–echo delay and other information-bearing features of echoes. These response properties are abundant in the auditory midbrain, thalamus, and cortex of the mustached bat. Combination-sensitive neurons utilize facilitatory and inhibitory neural interactions to create selective responses to acoustic features. The different interactions (facilitatory and inhibitory) originate in different stages within the ascending auditory pathway. Inhibitory interactions arise mostly in the auditory brainstem, while facilitatory interactions are within the inferior colliculus, but also display the results of the inhibitory interactions originating at brainstem levels. These response properties are well suited to the analysis of sonar echoes, but similar combinatorial response properties are tuned to spectral elements outside sonar frequency bands, probably to elements of social vocalizations. It is therefore likely that the neural analysis of biosonar signals shares common features with the analysis of other vocal signals. [Work supported by the National Institute on Deafness and Other Communication Disorders.]

3:55–4:10 Break

4:10

1pAB6. Pinniped bioacoustics: Atmospheric and hydrospheric signal production, reception, and function. Ronald J. Schusterman, David Kastak, Colleen Reichmuth Kastak, Marla Holt (Long Marine Lab., Univ. of California, 100 Shaffer Rd., Santa Cruz, CA 95060, rjchust@ucsc.edu), and Brandon L. Southall (Office of Protected Resources, Silver Spring, MD 20910)

There is no convincing evidence that any of the 33 pinniped species evolved acoustic specializations for echolocation. However, all species produce and localize signals amphibiously in different communicative contexts. In the setting of sexual selection, aquatic mating male phocids and walruses tend to emit underwater calls, while male otariids and phocids that breed terrestrially emit airborne calls. Signature vocalizations are widespread among pinnipeds. There is evidence that males use signature threat calls, and it is possible that vocal recognition may be used by territorial males to form categories consisting of neighbors and strangers. In terms of
mother–offspring recognition, both otariid females and their pups use acoustical cues for mutual recognition. In contrast, reunions between phocid females and their dependent pups depend mostly on pup vocalizations. In terms of signal reception, audiometric studies show that otariids are highly sensitive to aerial sounds but slightly less sensitive to underwater sounds. Conversely, except for deep-diving elephant seals, phocids are quite sensitive to acoustic signals both in air and under water. Finally, despite differences in absolute hearing sensitivity, pinnipeds have similar masked hearing capabilities in both media, supporting the notion that cochlear mechanics determine the effects of noise on hearing.

4:30

1pAB7. Cognitive processes in bird song. Jeffrey Cynx (Dept. of Psych., Vassar College, Poughkeepsie, NY 12604, chaos@vassar.edu)

Anthropomorphic hypotheses can alter previous ethological concepts. Songbirds have been traditionally categorized as open- or close-ended learners. Open-ended learners such as canaries and starlings continue to learn new songs throughout life. Close-ended learners such as song sparrows and zebra finches appear to learn song once and then repeat this song in a stereotyped or crystallized manner for the rest of their lives. Research over the last dozen years or so has produced evidence that whatever is close-ended in songbirds may be more than a little ajar. It is clear that adult song is a highly dynamic and closely monitored act. In these regards, it has a number of cognitive processes similar to human speech. Birds appear to continually monitor their own song, being able to stop in midsong if necessary. They also regulate the song amplitude given environmental and social conditions, and show song perturbations when experiencing delayed auditory feedback. However, so far as is known, close-ended learners cannot learn new song elements from a model, although there are hints to the contrary, including both behavioral and physiological results.

Contributed Papers

4:50

1pAB8. Human listening studies reveal insights into object features extracted by echolocating dolphins. Caroline M. DeLong (New College of Florida, 5700 N. Tamiami Trail, Sarasota, FL 34243, cdelong@ncf.edu), Whitlow W. L. Au (Hawaii Inst. of Marine Biol., Kailua, HI 96734), and Herbert L. Roitblat (DolphinSearch Inc., Ventura, CA 93001)

Echolocating dolphins extract object feature information from the acoustic parameters of object echoes. However, little is known about which object features are salient to dolphins or how they extract those features. To gain insight into how dolphins might be extracting feature information, human listeners were presented with echoes from objects used in a dolphin echoic-visual cross-modal matching task. Human participants performed a task similar to the one the dolphin had performed; however, echoic samples consisting of 23-echo trains were presented via headphones. The participants listened to the echoic sample and then visually selected the correct object from among three alternatives. The participants performed as well as or better than the dolphin (M = 88.0% correct), and reported using a combination of acoustic cues to extract object features (e.g., loudness, pitch, timbre). Participants frequently reported using the pattern of aural changes in the echoes across the echo train to identify the shape and structure of the objects (e.g., peaks in loudness or pitch). It is likely that dolphins also attend to the pattern of changes across echoes as objects are echolocated from different angles.

5:05


Echolocating animals like bats and toothed whales navigate and locate food by means of echoes from sounds transmitted by the animals themselves. Toothed whale echolocation has been studied intensively in captivity, but little information exists on how echolocation is used by wild animals for orientation and prey location. To expand on this issue, a noninvasive, acoustic Dtag (96-kHz sampling, 16-bit resolution) was deployed on two Blainvilles beaked whales. The tagged whales only clicked at depths below 200 m during deep foraging dives. The echolocation clicks are directional, 250-ms transients with peak energy in the 30–40-kHz band. Echoes from the seafloor and from prey items were recorded. The regular click rate is not adjusted to the decreasing echo delay from incoming prey until the target is within an approximate body length of the whale after which the click rate is increased rapidly akin to the buzz phase of echolocating bats. This suggests that the whales use different sonar strategies for operating in near versus far field modes. Changes in received echo intensities from prey targets during approaches are compared to the active gain control in the receiving system of bats and in the transmitting system of dolphins.

5:20

1pAB10. Aerial hearing sensitivity in some pinnipeds is comparable to that of humans. Colleen Reichmuth Kastak, David Kastak, Marla M. Holt, Ronald J. Schusterman (Univ. of California, Long Marine Lab., 100 Shaffer Rd., Santa Cruz, CA 95060, coll@usc.edu), and Brandon L. Southall (NOAA Fisheries Acoust. Prog., Silver Spring, MD 20910)

Aerial hearing sensitivity was measured in pinnipeds in a sound attenuating, hemi-anechoic chamber. Thresholds at 12 frequencies between 0.1 and 32.5 kHz were obtained behaviorally for three individuals (a California sea lion, a harbor seal, and a northern elephant seal) and compared to thresholds obtained using headphones in less controlled testing environments. The thresholds measured in the chamber revealed the expected relative changes in sensitivity with frequency; however, the absolute sensitivities were much better than had been previously measured. Harbor seal thresholds were on average 25 dB lower with best sensitivity of 2 dB (ref: 20 μPa) at 3.2 kHz. Elephant seal thresholds averaged 23 dB lower with best sensitivity of 27 dB (ref: 20 μPa) at 0.4 kHz. Thresholds for the California sea lion were also much lower than expected, with best sensitivity of 1 dB (ref: 20 μPa) at 12 kHz. The thresholds measured for the sea lion and harbor seal rival those of human subjects at some frequencies, and suggest that previously reported aerial hearing thresholds in pinnipeds were significantly noise limited. Further, the results indicate that these pinnipeds have greater sensitivity in air than in water when comparisons are made in terms of sound pressure.

5:35

1pAB11. A computational model of echolocation: Transformation of spectragram into the reflected intensity distribution for range discrimination of multiple closely spaced objects. Isaac Matsuo, Kenji Kunugiyama, and Masafumi Yano (Res. Inst. of Elec. Commun., Tohoku Univ., 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan, matsuo@riec.tohoku.ac.jp)

Using frequency-modulated echolocation, bats can discriminate the range of objects with an accuracy of less than a millimeter. However, the echolocation mechanism is not well understood. The delay separation of three or more closely spaced objects can be determined through analysis of the echo spectrum. However, delay times cannot be properly correlated with objects using only the echo spectrum because the sequence of delay
separations cannot be determined without information on temporal changes in the interference pattern of the echoes. To illustrate this, Gaussian chirplets with a carrier frequency compatible with bat emission sweep rates were used. The delay time for object 1, $T_1$, can be estimated from the echo spectrum around the onset time. The delay time for object 2 is obtained by adding $T_1$ to the delay separation between objects 1 and 2 (extracted from the first appearance of interference effects). Further objects can be located in sequence by this same procedure. This model can determine delay times for multiple closely spaced objects with an accuracy of about 1 microsecond, when all the objects are located within 30 microseconds of delay separation. This accuracy is possible even with objects having different reflected intensities and in a noisy environment.

MONDAY AFTERNOON, 24 MAY 2004

NEW YORK BALLROOM A, 1:00 TO 5:00 P.M.

Session 1pAO

Acoustical Oceanography and Underwater Acoustics: Geoacoustic Inversion

Kyle M. Becker, Chair

*Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804-0030*

**Contributed Papers**

1:00

1pAO1. Perturbative inversion method for range-varying seabed sound speed profile estimation. Luiz L. Souza (MIT/WHOI Joint Prog. in Oceanogr./Appl. Ocean Sci. and Eng., 77 Massachusetts Ave., Rm. 5-435; Cambridge, MA 02139) and George V. Frisk (Florida Atlantic Univ., Dania Beach, FL 33004)

In a range-dependent, slowly varying medium, the acoustic field can be represented as a sum of modes that adapt to the local properties of the medium. Modal wavenumbers measured as a function of source-receiver range, as described, for example, by Souza and Frisk (J. Acoust. Soc. Am. 113, 2204 (2003)), can be used to estimate seabed geoaoustic properties. An algorithm for estimating the compressional sound speed profile as a function of range is described. The technique consists of solving a Fredholm integral equation relating the perturbation of a background sound-speed profile to the perturbation of the modal eigenvalues [Rajan et al., J. Acoust. Soc. Am. 82, 998–1017 (1987)] in conjunction with smoothing constraints based on the stochastic regularization method. The formulation leads to a Kalman filter based algorithm, where the sound speed profile is obtained by solving the perturbative equations sequentially in range. The technique is applied to synthetic and experimental data from the modal mapping experiments (MOMAX).

1:15


A shallow water ocean acoustics experiment is described from which it is possible to determine in principle the scattering data necessary to recover the sound speed in the ocean bottom. This data is a “reflection coefficient.” The essential assumptions are that the speed in the ocean is known but unknown in the bottom and that the measurements are made in the ocean layer.

1:30

1pAO3. Simple single path calculations for MFP tomography. Alexandra Tolstoy (ATolstoy Sci., 8610 Battailles Court, Annandale, VA 22003, atolstoy@ieee.org)

MFP tomography promises a strong computational advantage over other geoacoustic inversion methods in 3-D, variable environments. In particular, the method can simply and rapidly estimate bottom parameters in full 3-D given average individual path estimates of the unknown properties. The method can then perform a simple linear inversion to obtain the full volume values. But, how simple is it to estimate the individual path averages? Sometimes simple, sometimes less so. This presentation will explore this issue.

1:45

1pAO4. Tabu evaluation in geoacoustic inversion. Zoi-Heleni Michalopoulou (Dept. of Mathematical Sci., New Jersey Inst. of Technol., Newark, NJ 07102, michalop@njit.edu)

Tabu optimization has recently been proposed as a tool for geoacoustic inversion. Tabu introduces a structured use of memory in the exploration of the search space, a novel element in inversion. For performance evaluation the technique is applied to Workshop 97 test cases. These cases provide a good framework for the validation of the new approach, since results from the application of several optimization schemes to these synthetic data sets exist in the literature. Tabu is evaluated on narrow-band and broadband data in the estimation of various combinations of geoaoustic parameters. The method is successful in accurate parameter estimation and requires relatively few forward model calculations. Tabu is also evaluated in a rotated coordinate framework, in an attempt to exploit parameter correlations for faster identification of the global maximum. [Work supported by ONR.]

2:00

1pAO5. Geoacoustic inversion of M-sequence data from experiments in the South Florida Straits. Ross Chapman and Yongmin Jiang (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada)

Matched field inversion is applied to low frequency data from experiments in the South Florida Straits to estimate a sediment velocity profile. The acoustic data consist of signal envelopes derived from M-sequences that were transmitted to a vertical hydrophone array over a range of 10 km. The sound speed profile supported a strong waveguide in the deeper parts of the water column. Signal propagation for a source in the waveguide was dominated by waveguide refracted paths, and surface-reflected bottom-reflected paths that interacted eight to ten times with the bottom. The very long range, shallow water geometry presents a significant challenge for an inversion based on matched field processing. However, examination of the spatial coherence of the received signal indicated useful spatial phase information for spectral components with high signal-to-noise ratios. A Gibbs sampling approach was used to test two different geoacoustic models for the site: (a) a constant gradient sediment and (b) a...
constant velocity sediment over a half space. The inversion showed a preference for slower speeds of around 1550 m/s at the sea floor, increasing to higher values up to 1700 m/s within about 100 m deeper. These values are consistent with ground truth at the site.

2:15


Array processing techniques in shallow water often require knowledge of the spatial coherence of sound propagation. However, few wideband data sets on this subject have been published. As a part of the ASIAEX program in the East China Sea, wideband sound propagation measurements were conducted in June 2001 along three radial directions and along a circle with a radius of 30 km. An earlier related paper dealt with reverberation data from this experiment [J. X. Zhou and X. Z. Zhang, J. Acoust. Soc. Am. 113, 2204 (2003)]. This paper analyzes sound propagation vertical coherence (SPVC) as a function of frequency, range, and propagation direction. Then the SPVC is used to invert bottom sound speed and attenuation. The results show that the characteristics of the SPVC and the SPVC-inverted bottom acoustic parameters agree well with geological survey and sediment cores in the ASIAEX area. [Work supported by ONR and NNSF of China.]

2:30

IpAO7. Geoacoustic inversion of broad-band ambient noise data using undersampled and short aperture arrays. Martin Siderius, Michael Porter (SAIC, Ctr. for Ocean Res., 10260 Campus Point Dr., San Diego, CA 92121), Chris Harrison (SACLANT Undersea Res. Ctr., 19138 La Spezia (SP), Italy), and The Kauai Group (SAIC, SPAWAR, UDel, APL-UW, MPL-UCSD, SSI, UNH, USM)

Ocean ambient noise is generated in many ways such as from winds, rain, and shipping. A technique has recently been developed [Harrison and Simons, J. Acoust. Soc. Am. 112 (2002)] that uses the vertical directionality of ambient noise to determine seabed properties. The ratio of beams steered towards the surface to those steered towards the bottom produces the bottom reflection loss curve. This technique was applied to data in the 200-1500 Hz band using a 16-m array. Extending this to higher frequencies allows the array length to be substantially shortened and greatly reduces interference from shipping. However, this limits the low end of the frequency spectrum since reduced aperture increases beam widths and the up/down beam ratio no longer produces reflectance loss. Similarly, for high frequencies, if hydrophone spacing is greater than half-wavelength, the beamformed output is aliased and the up/down ratio produces erroneous results. In general, frequencies much below the array design will suffer from large beams and frequencies above from undersampling. In this paper, we describe techniques for obtaining seabed properties from ambient noise data. The results have been presented from the KauaiEx (July 2003) and ElbaEx (October 2003) experiments.

2:45

IpAO8. Appraisal of an energy flux model for geoacoustic inversion of ambient noise coherence data. David J. Thomson and Francine Desharnais (Defence R&D Canada—Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada, david.thomson@drdc-rddc.gc.ca)

The vertical directionality of the ambient noise field in a shallow-water waveguide is highly dependent on the local geoacoustic properties of the seabed and is directly related to the vertical noise coherence observed between two closely spaced hydrophones. As a result, a direct measurement of the broadband ambient noise coherence can be used to invert for sea-bottom properties that affect propagation. In previous work, an energy flux model [D. M. F. Chapman, Proc. Inst. Acoust. 9(4), 1–11 (1987)] for computing vertical noise coherence in shallow water was extended to include the effects of multilayered geoacoustic seabeds, refraction and absorption within the water column, sensor-pair tilt, and nondipolar radiation patterns due to near-surface but finite source depths. This modified flux-based coherence model was combined with a hybrid local/global nonlinear optimization scheme and used to estimate geoacoustic source/sensor parameters for several synthetic data sets [F. Desharnais et al., J. Acoust. Soc. Am. 113, 2204 (2003)]. In this paper, the capability of the energy flux coherence model/geoacoustic inversion procedure is assessed for ambient noise data that were measured at several shallow-water sites over differing seabed types. Where possible, comparisons with geoacoustic properties estimated by other methods will be presented.

3:00–3:15 Break

3:15

IpAO9. Geoacoustic inversion from moving ship of opportunity in deep-water environment. David P. Knobles, Tim Scoggins, and Jack Shooter (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029)

This study examines the estimation of seabed geoacoustic parameters in a deep-water environment using acoustic data generated from the noise of a moving surface ship. The received sound field was recorded on a vertical line array in the North Pacific in approximately 5000 m of water. These data, which were recorded in 1973 in an experiment called Church Opal, have recently been re-digitized and made available to the underwater acoustics community. Spectrograms of time and frequency for hydrophones below the critical depth illustrate striation patterns associated with the inverse of the time interval between direct and bottom reflected rays paths. The cross-correlation of signals between phones on the VLA is thus sensitive to the structure of the seabed. A cost function of the difference between cross-phonospectra of modeled and measured received acoustic signals in the 10–600 Hz band is minimized using a simulated annealing method. Parameters that are estimated include those associated with the ship track and those associated with the seabed. Uncertainties of the various parameters are examined to determine the uniqueness of the various parameters estimated from the acoustic data. The estimated geoacoustic structure is compared to that reported in the literature. [Work supported by ONR.]

3:30

IpAO10. Iterative geoacoustic inversion. Peter Gerstoft, Chen-Fen Huang, and William S. Hodgkiss (Scripps Inst. of Oceanogr., La Jolla, CA 92039-0238)

In geoacoustic inversion we commonly invert each data block independently. But when these inversions are close in space and time it is often beneficial to use the results of the previous inversion as a starting condition for the next inversion. A natural framework for this is a Bayesian approach where the posterior inversion becomes the prior for the current inversion. It is implemented using a Metropolis–Hastings sampler. Such an approach potentially will allow us to work with lower SNR ratios, as in for example self-noise inversion. This approach has similarities to data assimilation and could lead to nowcasting of the acoustic field. [Work supported by ONR.]

3:45

IpAO11. Data error estimation for matched-field geoacoustic inversion. Michael J. Wilmut, Stan E. Dosso, and Jan Dettmer (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, mwjwilmut@uvic.ca)

Nonlinear Bayesian methods have been applied to geoacoustic inversion to estimate uncertainties for seabed parameters by sampling the posterior probability density. This procedure requires quantifying the errors on the acoustic data, including both measurement and theory errors, which are generally not well known. To date, point estimates for data errors have been derived using a global maximum likelihood approach. However, this is not consistent with the Bayesian formulation, and ignores the effects of uncertainty in the error estimates and interdependencies between the data.

errors and geoacoustic parameters. The Bayesian approach treats the data errors as random variables and includes them as additional parameters within the inversion. However, this increases significantly the number of unknowns and the computational effort. A third approach is to use a local maximum likelihood error estimate evaluated independently for each geoacoustic model considered in the sampling procedure. This has the benefit of not increasing the number of unknowns or computational effort, but includes some of the effects of the error uncertainties and interdependencies. The three approaches are compared for Bayesian matched-field geoacoustic inversion of both synthetic and experimental data.

4:00
1pAO12. Influence of data uncertainty on matched-field geoacoustic inversion. Chen-Fen Huang, Peter Gerstoft, and William S. Hodgkiss (Scripps Inst. of Oceanogr., La Jolla, CA 92037-0238)

Quantifying uncertainty in matched-field geoacoustic inversion using a Bayesian approach has attracted the attention of several authors in recent years. The complete solution to the inverse problem is given by the posterior probability distribution (PPD) in the model space given good knowledge of the noise statistics. Therefore, an estimate of noise (including measurement errors and model mismatch) is essential to obtain the PPDs. In this study, several approaches to accommodating the effects of the noise are suggested, such as using different likelihood functions or treating the noise as a nuisance parameter. The posterior probability of the model parameters is sampled using a Gibbs sampler approach based on the Metropolis algorithm. Comparisons between the approaches using synthetic data and experimental data are presented. [Work supported by ONR.]

4:15

Whereas bounding of errors in inverse problems is not generally possible, the problem has been shown to be tractable if the forward equation can be linearized about some (preferably the true) set of values and if, further, all errors are assumed Gaussian. Under these assumptions, the covariances of the a posteriori errors can be formulated, thus providing bounds on the uncertainty resulting from the inverse process. A previous effort analyzed the errors involved in benchmark cases and compared results to other published analyses. This surprisingly demonstrated the ability of a linearized relation to yield estimates of errors comparable to those achieved in Monte Carlo sampling. However, such a linear model neglects the interaction between model parameters. This current effort extends the analysis to second order. First and second derivatives of the acoustic field with respect to the environmental variables are calculated utilizing a parabolic equation propagation model. [Work supported by ONR.]

4:30
1pAO14. An equivalent transform method for evaluating the effect of water column mismatch on geoacoustic inversion. Ying-Tsong Lin, Chi-Fang Chen (Nat. Taiwan Univ., No. 1, Section 4, Roosevelt Rd., Taipei 106, Taiwan, yt.lin@msa.hinet.net), and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

An equivalent transform method for evaluating the effect of water column mismatch on geoacoustic inversion is presented. This quantitative method is derived from perturbative inverse formulations and gives a quantitative way to evaluate how the water column mismatch error transfers into the bottom inversion results. In this paper, this method is illustrated with two test cases, in which linear and nonlinear internal waves are considered as the cause of the water column mismatch, respectively. In the first case, the range-averaged geoacoustic inversion errors due to the water column mismatch are largely eliminated after a full cycle of the linear internal wave, and the error pattern repeats with the period of the internal wave. However, in the second case, the errors are accumulated but scaled down with increasing range. Additionally, ignoring the lower-frequency components of the signal removes its corresponding contribution from the overall errors, and reduces the bottom depth influenced by water column mismatch.

4:45

The effects of environmental variability on acoustic propagation and sonar performance in shallow water has received much attention in recent literature. Of particular interest is the influence of water column variability on acoustic propagation and the subsequent effects on geoacoustic inversion. In this paper, a modal-based inversion method is considered with the inclusion of internal waves in a shallow-water waveguide. The input data to the inversion algorithm are estimates of horizontal wavenumbers corresponding to the propagating modes of a cw point-source field. Often, wavenumber estimation is performed under the assumption that the pressure field can be represented by an adiabatic mode sum. However, it is well known that sound speed fluctuations in the water column due to internal waves promote the coupling of energy between propagating acoustic modes, thus violating the adiabatic assumption. Nevertheless, it was recently demonstrated that sound speed fluctuations in the water column due to a weak deterministic internal wave field can enhance spectral wavenumber estimates [Becker and Frisk, in Impact of Littoral Variability on Acoustic Predictions and Sonar Performance (Kluwer, Dordrecht, 2002)]. That work is extended to compare inversion results for wavenumber estimates obtained both with and without internal waves. [Work supported by ONR Ocean Acoustics.]


**1pBB1. Sonoelastography imaging: Principles and practices.** Lawrence S. Taylor (BME Dept., Univ. of Rochester, 310 Hopeman Hall, Rochester, NY 14607, lstaylor@bme.rochester.edu), Deborah J. Rubens, and Kevin J. Parker (Univ. of Rochester, Rochester, NY 14607)

Vibration amplitude sonoelastography imaging is an ultrasound imaging technique in which low frequency (100–500 Hz), low amplitude shear waves are propagated deep into tissue, while real time color Doppler techniques are used to image the resulting vibration field. A radio frequency ultrasound signal is phase modulated by a vibrating particle such that the peak vibration amplitude is directly proportional to the spread (standard deviation) of its power spectrum. A key application for this technique is the detection and imaging of small lesions. Finite element studies predict that a discrete hard inhomogeneity present within a larger region of soft tissue will cause a decrease in the vibration field at the location of the inhomogeneity. The inhomogeneity is made visible as a region of low vibration in the color Doppler image. The principles and practices of this technique are reviewed and results are shown for the detection and imaging of stiff thermal lesions induced in bovine calf liver.

**1pBB2. Recent advances in elastography.** Jonathan Ophir (Univ. of Texas Med. School, Ultrason. Lab., 6431 Fannin, Houston, TX 77030, jophir@uth.tmc.edu)

No known modality is capable of imaging the elastic properties of tissue directly. It is therefore necessary to apply an external or internal mechanical stimulus to the tissue system, and to observe the tissue response in terms of local internal deformations. The use of ultrasound for this purpose has several important advantages that include real-time imaging, high resolution in motion estimation (1 micron), simplicity, noninvasiveness, and low cost. Elastography is performed by obtaining a set of ultrasonic echo signals from a target, subjecting the target to a small axial deformation and obtaining a second set of echo signals. Time-delay estimations along the direction of the applied load are computed. Using the gradient operator, the time-shift estimates are then converted to strain information, which is displayed in the form of a two-dimensional image named elastograms. We demonstrate that the strain distributions in tissues are well correlated with the distribution of tissue elastic moduli under certain conditions. It is also possible to obtain lateral strain elastograms and Poissons ratio elastograms, which convey additional tissue mechanical information. In this presentation we will give an overview of elastography and describe some of the newer capabilities and clinical applications. [Work supported by NIH Grant P01CA64597.]


Interstitially implanted radioactive seeds are becoming a popular and effective means of treating a wide range of diseases, primarily cancer. Brachytherapy of prostate cancer, for example, is proving to be as effective, in terms of 5-year survival, as surgery for management of gland-confined disease, and to many, the side effects of brachytherapy are less debilitating than those of surgery. While seeds are implanted under guidance from transrectal ultrasound, gland movement and distortion during implantation may lead to seed misplacement; implanted seeds are difficult to visualize ultrasonically after implantation for a variety of reasons. We are investigating a variety of ultrasonic techniques intended to overcome the limitations of conventional ultrasonic imaging, and the following three methods are showing encouraging promise: resonance/vibration imaging; modified elastography; and two-dimensional correlation. The resonance/vibration method induces natural vibrations in seeds and, using power-Doppler methods, images the movement these vibrations induce in immediately adjacent tissues. The modified-elastography method uses standard elastographic processing with higher-than-normal applied strains so that soft-tissue signals decorrelate while echo signals from rigid seeds retain correlation. The two-dimensional correlation method compares echo signals from tissue with those from a reference seed and shows seed locations by mapping the resulting correlation coefficients.

Acoustic radiation force impulse (ARFI) imaging utilizes brief, high energy, focused acoustic pulses to generate radiation force in tissue, and conventional diagnostic ultrasound methods to detect the resulting tissue displacements in order to image the relative mechanical properties of tissue. The magnitude and spatial extent of the applied force is dependent upon the transmit beam parameters and the tissue attenuation. Forcing volumes are on the order of 5 mm³, pulse durations are less than 1 ms, and tissue displacements are typically several microns. Images of tissue displacement reflect local tissue stiffness, with softer tissues (e.g., fat) displacing farther than stiffer tissues (e.g., muscle). Parametric images of maximum displacement, time to peak displacement, and recovery time provide information about tissue material properties and structure. In both in vivo and ex vivo data, structures shown in matched B-mode images are in good agreement with those shown in ARFI images, with comparable resolution. Potential clinical applications under investigation include soft tissue lesion characterization, assessment of focal atherosclerosis, and imaging of thermal lesion formation during tissue ablation procedures. Results from ongoing studies will be presented. [Work supported by NIH Grant R01 EB002132-03, and the Whitaker Foundation. System support from Siemens Medical Solutions USA, Inc.]

1pBB5. Elasticity imaging of arterial wall with transcutaneous ultrasound both in longitudinal-axis and short-axis planes. Hiroshi Kanai and Hideyuki Hasegawa (Dept. of Electron. Eng., Tohoku Univ., Aramaki-Aza-Aoba 05, Sendai 980-8579, Japan)

A method for measuring regional elasticity of tissue surrounding atherosclerotic plaque is described. An ultrasonic beam was scanned with a conventional linear-type probe, and multiple layers were preset from luminal surface to adventitia of the common carotid artery (CCA) with intervals of 375 μm. By applying the method [IEEE Trans. UFFC 46, 1229–1241 (1999)], a minute decrease of several tenths of a micrometer in thickness of each layer resulting from arrival of the pressure wave was determined. By assuming that the arterial wall is incompressible and that the blood pressure is applied normal to each layer, the elastic modulus in the circumferential direction of each layer was estimated at intervals of 75 μm in the radial direction and 150–300 μm in longitudinal direction. On the other hand, by designing the directions of ultrasonic beams so that each beam always passes through the center of the artery, the cross-sectional elasticity image in the short-axis plane was also obtained. Based on the elasticity library determined by comparing the elasticity distribution and their pathological images, each point was statistically categorized as lipid, a mixture of smooth muscle and collagen fiber, or other. By applying the method to the CCAs, soft inclusion of lipid was found for plaques.

1pBB6. Synergy of ultrasound, elasticity, and optoacoustic imaging for improved detection and differentiation of cancerous tissue. Stanislav Emelianov, Salavat Aglyamov, Jignesh Shah, Shriram Sethuraman (Dept. of Biomed. Eng., Univ. of Texas, Austin, TX 78712, emelian@mail.utexas.edu), Guy Scott, Rainer Schmitt (WinProbe Corp., North Palm Beach, FL 33408), Andrei Karpionou, Massoud Motamedi (Univ. of Texas Medical Branch, Galveston, TX 77555), and Alexander Oraevsky (Fairway Medical Technologies, Inc., Houston, TX 77099)

The effective management of cancer requires early yet reliable detection, localization, and diagnosis. Therefore, there is a definite and urgent clinical need for an imaging technique that is widely available, is simple to perform, is safe, and that can detect and adequately diagnose cancer. In this paper we present a hybrid imaging technology based on fusion of complementary imaging modalities ultrasound, optoacoustics, and elastography to take full advantage of the many synergistic features of these modalities and thus to significantly improve sensitivity and specificity of cancer imaging. To evaluate our approach, numerical and experimental studies were performed using heterogeneous phantoms where ultrasound, optical, and viscoelastic properties of the materials were chosen to closely mimic soft tissue. The results of this study suggest that combined ultrasound-based imaging is possible and can provide more accurate, reliable and earlier detection and diagnosis of tissue pathology. In addition, monitoring of cancer treatment and guidance of tissue biopsy are possible with combined imaging system. Practical and experimental aspects of combined imaging will be discussed with emphasis on data capture and signal/image processing algorithms. The paper will conclude with a discussion of the advantages, limitations, and potential clinical applications of the combined imaging technique.

Contributed Papers

3:30

3:30

1pBB7. Studying viscoelasticity in soft tissues with supersonic shear imaging. Jeremy Bercoff, Marie Muller, Mickael Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, 10 rue Vauquelin, 75005 Paris, France, jeremy.bercoff@lsa.espci.fr)

In this work, a 3D analytical formulation of the mechanical Green’s function in a viscoelastic medium is derived and presented. Based on a Voigt model to take into account viscoelasticity, this mathematical formulation is validated experimentally using the supersonic shear imaging technique (SSI). Taking benefit of the ultrasonic remote generation of a moving shear source radiating low-frequency shear waves in the medium, this technique has been studied and validated for soft tissue elasticity mapping in previous works. It is shown here that the spatial and temporal shape of shear waves induced in soft tissues using SSI can be accurately modeled with the viscoelastic Green’s function. The influences of important parameters such as viscosity, elasticity, or diffraction on the shear wave shape are carefully studied and discriminated. In a second part, taking advantage of the previous modeling, the inverse problem consisting of recovering shear elasticity and viscosity is presented and validated using the Green’s function-based simulation tool. Experiments on tissue-mimicking phan-
tompresenting different viscoelastic properties are presented. The influence of out-of-plane shear propagation on the inversion algorithm is discussed.

3:45
1pBB8. A simple viscoelastic model for soft tissues in the frequency range 5–15 MHz. Ximai Yang and Charles Church (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, xmyang@olemiss.edu)

Measurements of the shear properties of soft tissues have been performed using a simple pulse–echo reflection technique. Briefly, a shear wave, generated by a shear wave transducer, propagates through a quartz rod, is reflected at the quartz–tissue interface, backpropagates through the rod, and is received by the transducer. The phase shift and the change in the magnitude of the wave with and without a tissue sample in place are measured, and the shear mechanical impedances are calculated and the complex modulus is determined. A negative shear modulus is often obtained in this frequency range, and none of the traditional linear models can explain this result. We propose a simple extension to the Voigt model for viscoelasticity to account for the effects of mass. In this model, a negative storage modulus is possible under the conditions of high frequency and long relaxation time. When the frequency is low and relaxation time is small, the new model reduces to the classic Voigt model. Based on this model, we will present the results of measurements of shear modulus, shear viscosity, and the constant term related to mass for several soft tissues. [Work supported by the US Army.]

4:00
1pBB9. Thermal ablation monitoring using radiation force-induced steady-state tissue motion. Hesheng Wang, Yun Zhou, and Cheri Deng (Dept. of Biomed. Eng., Case Western Reserve Univ., 10900 Euclid Ave., Cleveland, OH 44106, cxd54@cwru.edu)

High-intensity focused ultrasound (HIFU) increases temperature in tissue and also induces local tissue movement during HIFU ablation. Such induced motion can be exploited to detect changes in tissue stiffness to monitor HIFU treatment. To implement real-time HIFU treatment monitoring, a HIFU duration is segmented into a sequence of on and off periods. The off periods are required for unaffected interrogation and are sufficiently short to minimize interruption to the intended thermal treatment. It is demonstrated that the displacements generated during the relatively long on period reach steady state and undergo recovery during the next off period. Our study focuses on the investigation of the steady-state displacement and its recovery to detect changes in tissue elasticity without additional push pulses during an HIFU treatment exposure. An integrated model and finite difference algorithm are developed to study how the dynamic displacements are related to HIFU characteristics and tissue properties. Our experiments utilized an ultrasound imaging system. The induced displacement and its recovery are estimated from the acquired backscattered signals using a cross-correlation algorithm. Optimal HIFU on and off periods are investigated. Our results indicate that the steady-state displacements characterize the changes in tissue elasticity accompanying lesion formation during HIFU ablation.

4:15

Using ultrasound focused beams, mechanical sources radiating low-frequency shear waves can be generated inside the body. Such waves propagate inside tissues at a few m/s and are sensitive to the mechanical properties of soft tissues. The ultrafast scanner developed at our laboratory (5000 images/s) is able to generate shear sources and image, in real time, the propagation of the resulting shear waves. Pushing and imaging sequences can be interleaved as desired. This great versatility enables various strategies concerning the pushing sequence. In particular, the super-sonic regime, which consists of moving the shear source at a speed higher than shear wave speed, has been introduced in previous works. In this work, a detailed vectorial analysis of the supersonic regime is presented using a 3D Green’s function formalism. By varying the speed and the trajectory of the shear source, conical waves of different angles, curved waves, or focused waves can be generated. Furthermore, changing the amplitude of the pushing sequences allows the temporal modulation of the resulting shear waves, leading to a complete control of the spatio-temporal shape of the shear wave. These palpation strategies are carefully studied using numerical simulations and tested on tissue-mimicking phantoms.

4:30
1pBB11. Noninvasive evaluation of complex elastic modulus of arterial vessels by ultrasound. Xiaoming Zhang, Mostafa Fatemi, and James F. Greenleaf (Dept. of Physiol. and Biomed. Eng., Mayo Clinic College of Medicine, Rochester, MN 55905)

Pulse wave velocity (PWV) is widely used for estimating the stiffness of an artery. It is well known that a stiffened artery can be associated with disease and aging. Usually, PWV is measured using the foot-to-foot time-delay method. However, the foot of the pulse wave cannot be accurately determined due to reflected waves. In addition, PWV is an average indicator of artery stiffness between the two measuring points. We propose to generate a wave in the artery wall using localized ultrasound radiation force. The wave velocity can be measured accurately by the phase change versus distance over a few millimeters. A mathematical model is developed by which the real part of the complex Young’s modulus of the artery can be determined from measured wave velocities. The imaginary part of the complex modulus can be assessed by the wave amplitude decay over distance. Experiments were conducted on a pig carotid artery in gelatin. The measured wave velocity is about 3 m/s at 100 Hz and 6.5 m/s at 500 Hz. The real part of complex modulus is estimated to be 300 KPa. The present technique offers a new method for estimating the complex elastic modulus of the artery over a short length.

4:45
1pBB12. Improvement in elastographic signal-to-noise ratio and resolution using coded excitation in elastography. Remi Souchon, Jean-Christophe Bera, Agnes Pousse, and Jean-Yves Chapelon (INSERM U556, 151 cours Albert Thomas, 69424 Lyon cedex 03, France, souchon@lyon.inserm.fr)

Coded excitation has been used in conventional sonography to improve the sonographic signal-to-noise ratio independently of resolution. In elastography, a tradeoff exists between spatial resolution and the elastographic signal-to-noise ratio (SNRe). In the present work, the use of coded excitation was investigated to remove this ambiguity in elastography. Both numerical simulations and phantom experiments were carried out to estimate the SNRe in a homogeneous material and the resolution in a material containing calibrated inclusions. These results were compared with those obtained using conventional pulse excitation. Various codes (Golay, Barker, chirp) and code lengths were tested. The numerical simulations used a simple 1D backscattering model to show the theoretical effects of coded excitation. Experiments were carried out using a more realistic setup based on a sector-scan imaging probe. In the absence of sonographic noise, simulations showed that codes induced only a slight decrease in SNRe at no cost in resolution. When sonographic noise was added into the model, a large improvement in SNRe was obtained at constant resolution. The experimental results corroborated these findings. [Work supported in part by National Cancer Institute (USA) Program Project Grant PO1-CA64597.]
Lindstaedt/H20849

reduced by lateral motion correction based on block-matching algorithms. This effect is caused by natural pulsatile blood flow. The strain imaging algorithm estimates radial strain from rf data based on frame-to-frame correlation. Rotating transducers often show nonuniform rotational distortion (NURD), which leads to misaligned echo lines in consecutive frames. This results in lateral motion artifacts and causes decorrelation. This effect is reduced by lateral motion correction based on block-matching algorithms. Results show that strain imaging can successfully be performed in vivo with data acquired predominantly in diastole. Different coronary tissue regions can be identified by local strain variations. If NURD is present, strain image quality is degraded. In some cases NURD is reduced by repositioning the transducer.

MONDAY AFTERNOON, 24 MAY 2004

Session IpNS

Noise: Noise Source Characterization

Richard H. Lyon, Chair
RH Lyon Corp, 691 Concord Avenue, Cambridge, Massachusetts 02138

Chair’s Introduction—1:25

Contributed Papers

1:30

IpNS1. The noise-cooling tradeoff in electronic equipment. Richard H. Lyon (RH Lyon Corp, 691 Concord Ave., Cambridge, MA 02138)

The noise produced by cooling air passing through electronics packages arises from two sources. One source is the noise of the air-moving fan of either an axial or centrifugal type. This noise may have tonal components and both those components and the broadband noise are dependent on the way that the fan is placed in the unit and on how close the operation is to the design operating point. Often, this can be the dominant noise source. The other source produces random noise due to the turbulent flow of air through the unit. Because the turbulent airflow is also responsible for heat transfer between the components and the air stream, we can regard this part of the noise as the irreducible noise due to cooling. If fan noise were eliminated, this part of the noise must remain. There is a relation therefore between the irreducible noise and the cooling of the unit. But the fan noise must also be considered. The relation between total airflow related noise and cooling requirements is developed in this paper.

1:45

IpNS2. Design and application of a perforator element muffer for reducing noise from fan-trays in telecommunication enclosures. Hugh Holness and M. G. Prasad (Dept. of Mech. Eng., Stevens Inst. of Technol., Hoboken, NJ 07030, hhlness@stevens.edu)

Generally the mufflers for reducing noise from fan-trays in telecommunication enclosures use sound absorptive materials in their design. Our objective in this work is to design a perforator element muffer that will produce similar acoustical performance without increasing pressure drop in the system as compared to a generally used muffer with sound absorptive material. The work includes parametric studies on the influence of perforated element on noise attenuation and pressure drop. The advantages and disadvantages of reactive perforator element mufflers over the dissipative mufflers in such applications will be studied. The studies include both design and experimental work.

2:00

IpNS3. Road traffic noise impact assessment based on the difference between existing and future traffic noise levels. Chetlur G. Balachandr, Sandor Juhasz, Arthur Morrone (Parsons Brinckerhoff, One Penn Plaza, New York, NY 10117, Balachandri@pbworld.com), and Noemi Castillo (HDR Inc., White Plains, NY 10004)

New York City has established standards for impact assessment based on the difference in traffic noise levels between the existing and future conditions. This difference is directly related to the logarithmic ratio of passenger car equivalents under the two conditions. The FHWA Traffic Noise Model 2.0 predicts noise levels from data based on the number of vehicles classified by category, average vehicle speed, and the distance to receptor. The model results can be used to determine the change in noise level between the existing and future conditions. Both approaches give essentially the same result. In heavily populated urban centers like New York City, where the measured baseline existing noise is not only from road traffic but also from many other noise sources such as air-

Downloaded 02 Feb 2012 to 131.174.248.116. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp
conditioning units, people noise, construction activities, etc., a new approach is needed to determine the change between the existing and future noise levels. This approach utilizes data obtained from field noise measurements which are conducted with and without trucks. This realistic procedure can be used to determine the number of new heavy trucks that could be introduced into the existing traffic stream without exceeding impact thresholds set by the New York City Noise Regulations.

2:15

IpNS4. Determination of sound power levels of some pyrotechnic devices using sound pressure measurements. Weixiong Wu (117 E. 29th St., New York, NY 10016)

A noise measurement study was conducted at five reservoirs in New York State during the winter and summer of 2003 to determine the sound power levels generated by some specific pyrotechnic devices (bangers, screamers, and CAPAs) using sound pressure techniques. The study was performed in support of an environmental impact statement (EIS) that defined the areas around each of these reservoirs where any significant impacts would occur as a result of the pyrotechnic operations. Sound characteristic data for each pyrotechnic device was collected at short distances with a direct line of sight to the pyrotechnic sources. The sound pressure levels for each pyrotechnic device were measured in two conditions with a direct line of sight to the pyrotechnic sources. The sound power levels, including both A-weighted and C-weighted 1/1 octave band values, were calculated based upon the measured sound pressure levels. With the absence of literature and manufacturing data supports, the values of this study were believed to be the most comprehensive emission data for the EIS noise analysis.

MONDAY AFTERNOON, 24 MAY 2004

NEW YORK BALLROOM B, 1:00 TO 4:45 P.M.

Session 1pPAa

Physical Acoustics: Recent Advances in Buried Landmine Detection II

James M. Sabatier, Cochair
University of Mississippi, National Center for Physics Acoustics, Coliseum Drive, University, Mississippi 38677

Bradley Libbey, Cochair
Night Vision and Electronic Sensors Directorate, 10221 Burbeck Road, Fort Belvoir, Virginia 22060

Invited Papers

1:00

1pPAa1. Recent developments in laser Doppler velocimeter for landmine detection. Amit Lal, Vyacheslav Aranchuk, and Cecil Hess (MetroLaser, 2572 White Rd., Irvine, CA 92614)

This presentation discusses laser Doppler vibrometer (LDV) instruments developed by MetroLaser that are specifically tailored to work with the acoustic/seismic technique for buried landmine detection and data obtained with these systems under various field conditions. The first system is a simple and rugged diode laser LDV that measures a single point on the ground. Multiple diode systems or a scanning mirror placed in front of the diode LDV are some of the strategies used to measure multiple points. The second system is a multi-beam LDV that uses a single YAG laser and diffraction optical elements to create an array of laser beams and thus multiple measurement points on the ground. The data obtained with LDV consist of velocity profiles at each point of the surface illuminated by each of the beams. By interrogating multiple points (with multiple vibrometers or with a multi-beam vibrometer) the system can ascertain with high probability the presence of a buried landmine. The systems exhibit a velocity resolution of about 1 micron per second, which is suitable for landmine detection. Low-frequency (100 kHz) carrier is employed in the multi-beam system to further enhance its performance with the acoustic/seismic technique.

2:30

IpNS5. Environmental noise impact of modern wind farms. Andrew Piacsek and Greg Wagner (Central Washington Univ., 400 E. University Way, Ellensburg, WA 98926-7422, piacsek@cwu.edu)

Electric power production from wind turbines has increased substantially during the past few years due to the growing emphasis on renewable energy sources and more efficient wind turbine technology. Although modern turbines are significantly quieter than early models, wind farms that are proposed near residential areas generate concern about potential noise issues. The present study consists of two parts: (1) the measurement of sound levels within a 2-km radius of the existing Nine Canyons wind farm near Richland, WA, and (2) the application of an outdoor sound propagation model to predict noise levels, both at Nine Canyons and in the vicinity of a proposed farm near Ellensburg, WA. At most locations within the Nine Canyon site, recorded sound levels were less than predicted levels, with the exception of some downwind sites that were lower in elevation than the source. Noise levels were greatest downwind from the turbines, but never exceeded 50 dBA beyond 500 m from the nearest turbine. In many cases, wind noise at the microphone exceeded noise levels from the turbines.

2:45

IpNS6. Acoustic project for installation of motor generator group by means of computer simulation. Jose C. Ferreira and Paulo T. Zannin (Lab Acustica Ambiental, Universidade Federal do Parana, Dept. Engenharia Mecanica, Centro Politecnico)

This work presents an acoustical project for the installation of a motor generator group of electricity in a hotel by means of computer modeling. The noise levels at the site have been obtained without the motor generator group, and via the computer modeling it has been deduced how these levels would be after the installation of the equipment. A possible solution to mitigate the noise impact the equipment would cause on the neighborhood has been indicated, and it has been predicted how the impact would be reduced after the implantation of this solution.
1:20

1pPAa2. Investigation of a combined seismic, radar, and induction sensor for landmine detection. Waymond Scott, Jr., Kangwook Kim (School of Elec. and Computer Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0250, waymond.scott@ece.gatech.edu), and Gregg Larson (Georgia Inst. of Technol., Atlanta, GA 30332-0405)

An experimental system to collect co-located ground penetrating radar (GPR), electromagnetic induction (EMI), and seismic data was developed to investigate the benefits of the fusion of these sensors. In the experiments, a range of mines and clutter objects were buried at various depths in the sandbox at Georgia Tech. Multiple burial scenarios were investigated with a variety of antipersonnel and antitank mines and typical clutter objects. The seismic system used in these experiments is an extension of our existing seismic mine detection system. The system uses electrodynamic shakers to generate a seismic wave which propagates across the simulated minefield, and a specially designed radar is used to measure the displacement of the surface caused by the seismic wave. The GPR makes use of modified resistive-vee antennas and operates over the frequency range of 500 MHz to 8 GHz. These antennas are very clean in that they have very little self-clutter and very low radar cross section to lessen the reflections between the ground and the antennas. The EMI sensor collects broadband data (300 Hz to 100 kHz) so that the relaxation frequencies of the buried targets can be used to aid discrimination. [Work supported by ARO.]

1:40


The use of ground-penetrating radar (GPR) for landmine detection and its relationship to acoustic–seismic (A/S) techniques is discussed. While GPR is very effective against buried metal targets, modern plastic-based landmines are much more difficult for GPR to detect. This difficulty is due to the low dielectric contrast between the landmine material and the soil itself. Recent advances in radar technology and signal-processing algorithms have improved GPR performance against buried plastic landmines. Several multichannel and synthetic aperture GPR systems have been developed that are capable of acquiring broad spectrum data over densely sampled spatial grids. Volumetric images of the subsurface produced by these systems are processed using automated-target-recognition (ATR) routines in order to determine the presence/absence of a landmine target. While GPR systems are capable of performing well on their own, the fusion of GPR data with data from an additional “orthogonal” sensor greatly increases the overall probability of detection (Pd) and reduces the false-alarm rate (FAR). Acoustic–seismic (A/S) sensors provide this orthogonality by exploiting mechanical properties of the target as opposed to the electromagnetic properties detected by GPR.

2:00


Recent success in using a laser Doppler vibrometer (LDV) based acoustic-to-seismic landmine detection [J. M. Sabatier and N. Xiang, IEEE Trans. Geosci. Remote Sensing 39, 1146–1154 (2001); N. Xiang and J. M. Sabatier, J. Acoust. Soc. Am. 113, 1333–1341 (2003)] and a ground penetrating synthetic aperture radar (GPSAR) suggested a novel configuration of fused sensors comprised of a LDV-based A/S detection sensor and a GPSAR. Extensive field experiments revealed that these two technologies can be considered orthogonal. When used in concert, a fused configuration may significantly improve the probability of detection and reduce the false alarm rate. They function best against different types of landmines under different burial conditions because they exploit disparate phenomena to detect mines. In order to optimize the fused detection ability using the two modalities, co-located field experiments have been conducted using both LDV-based A/S sensor and a GPSAR sensor. This paper will discuss the comparative experimental study using the recent co-located field scanning results.

2:20

1pPAa5. Landmine detection from acoustic/seismic signals using ordered weight-averaging operators. Paul Gader, Roopnath Grandhi, Ali Hocaoglu (301 CSE Bldg., CISE Dept., Univ. of Florida, Gainesville, FL 32611, pgader@cise.ufl.edu), James Keller, and Tsaipei Wang (Univ. of Missouri--Columbia, Columbia, MO 65211)

Acoustic/seismic systems demonstrate significant potential for reliable detection of landmines, particularly when coupled with ground-penetrating radar (GPR). Acoustic/seismic systems provide complementary information to GPR. One difficulty in developing detection algorithms for acoustic/seismic systems is that the frequency bands at which information occurs vary according to factors such as mine depth and mechanical properties of the soil, which are not known during mine detection. Ordered weighted-averaging (OWA) systems offer the potential for robust detection of landmines under different conditions because they naturally provide the capability to process the best collection of frequency bands in a fashion that can be optimized. In this paper, a processing framework is presented that combines OWA operators for feature analysis with decision making. This framework can be optimized using acoustic/seismic signals as well as acoustic/seismic signals combined with GPR. In the latter case, the optimization can help to provide the appropriate mechanism for combining the complementary information from both sensors. Results are presented using real data from both mines and other buried objects collected at an outdoor test site.
Synchronized with the vibration peak and the vibration valley, respectively, vibration measurement of the ground surface. The advantage of parallel optical processing. An optical technique that lends itself to such vibration sensing is an electronic speckle pattern interferometer for buried landmines that cues on such modal patterns must take advantage of the velocity distribution functions. Furthermore, the velocity distribution can be determined from the width of the vibration amplitude versus frequency is measured on a swept spectrum analyzer using an accelerometer located on the soil-air interface and reference wavefront and recorded in two video frames by using a CCD camera. The spatial distribution of the 2D vibration amplitude is obtained by subtracting two speckle patterns and processing the corresponding fringe pattern. The displacement sensitivity of the ESPI is about 30 nm. Here, we not only consider airborne sound sources to excite vibrations in the ground, but also a mechanical shaker to significantly increase the vibration amplitudes at the spot of interest. The vibrational pattern at the surface of the ground over buried antitank and antipersonnel landmines is studied.

Contributed Papers

2:40
IpPAAa6. An acoustic landmine detection confirmatory sensor using continuously scanning multibeam LDV. James M. Sabatier (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677, sabatier@olemiss.edu), and Vyacheslav Aranchuk (MetroLaser, Inc., Irvine, CA 92614-6236)

Acoustic-to-seismic coupling technology using an LDV as a vibration sensor has proved itself as a potential confirmative sensor for buried landmine detection. One of the most important objectives of this technology is to increase the speed of measurements. Traditionally used point-by-point scanning LDV systems cannot provide fast measurements. A moving platform that uses 16 LDVs and a continuously scanning single-beam LDV has been used to increase the speed of detection of buried landmines. Recently a multibeam LDV simultaneously probing 16 positions on the ground has been developed and successfully used for landmine detection. In this work, we report the use of a continuously scanning multibeam LDV as a confirmatory sensor for acoustic landmine detection. The multibeam LDV simultaneously illuminates the ground in 16 points spread over a 1-m line. A scanning mirror moves all 16 laser beams across the line. An airborne sound source in the frequency range of 80–300 Hz has been used to excite vibrations in the ground. The system enables scanning a 1-m by 1-m area and provides the vibrational pattern at the surface of the ground in a much shorter time than with previous scanning techniques.

2:55
IpPAAa7. A different approach for processing multibeam laser Doppler vibrometer (LDV) data. Ronald A. Wagstaff and Kenneth E. Gilbert (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677, rwagstaff@olemiss.edu)

Exciting the ground with an acoustic tonal projected by a loudspeaker is a well-known method for detecting buried landmines. The subsequent ground motion is measured with a laser Doppler vibrometer (LDV). The LDV data contain the tonal in a frequency-modulated waveform. One approach for demodulating the data and extracting the tonal uses a Hilbert transform. The ground velocity can be obtained from these data to identify mine presence or absence. An alternate approach to mine detection is to perform consecutive fast Fourier transforms on the modulated LDV data, and to average the output powers in each spectral bin. This results in a ground velocity distribution function in the spectrum. The proximity of the beams to a mine (over, near, not near) can be determined from the width of the velocity distribution functions. Furthermore, the velocity distribution functions provide additional information that previous techniques do not. Such information may be useful, e.g., the possibility of separating mines from false targets. This new technique will be discussed, and the results from measured multibeam LDV data will be presented. [This material is based upon work supported by the U. S. Army Communications-Electronics Command Night Vision and Electronic Sensors Directorate under Contract DAAB15-02-C-0024.]

3:10–3:30 Break

3:30
IpPAAa8. Rapid high spatial resolution imaging of ground vibration for buried landmine detection using ESPI. William C. Alberts II, James M. Sabatier (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677, walberts@olemiss.edu), and Vyacheslav Aranchuk (MetroLaser, Inc., Irvine, CA 92614-6236)

Recent work has shown that many landmines exhibit a multimode vibration pattern. To fully map the vibration pattern of these modes requires spatial resolutions on the order of millimeters. Any practical field instrument to detect landmines that cues on such modal patterns must take advantage of parallel optical processing. An optical technique that lends itself to such vibration sensing is an electronic speckle pattern interferometer (ESPI). The double-pulsed ESPI system has been used for the vibration measurement of the ground surface. The first and second laser pulses synchronized with the vibration peak and the vibration valley, respectively, illuminate the object. The imaged object wavefronts are combined with a reference wavefront and recorded in two video frames by using a CCD camera. The spatial distribution of the 2D vibration amplitude is obtained by subtracting two speckle patterns and processing the corresponding fringe pattern. The displacement sensitivity of the ESPI is about 30 nm. Here, we not only consider airborne sound sources to excite vibrations in the ground, but also a mechanical shaker to significantly increase the vibration amplitudes at the spot of interest. The vibrational pattern at the surface of the ground over buried antitank and antipersonnel landmines is studied.

3:45

An ultrasonic displacement sensor is currently being investigated. A possible application for this sensor is acoustic/seismic landmine detection in which a noncontact vibrometer is employed to measure the normal surface velocity of soil. The ultrasonic sensor is an alternative displacement sensor to a laser Doppler vibrometer (LDV). The ultrasonic wavelength suggests that the ultrasonic sensor is capable of resolving displacements in the presence of rough surfaces. This functionality is of particular interest because scattering of optical signals by rough surfaces can prohibit accurate LDV measurements. The ultrasonic sensor will be evaluated in an array of surface roughness conditions. The performance of the ultrasonic sensor will be compared to the LDV for each of these cases. The relative advantages and limitations of the ultrasound system will be presented. The applicability of the ultrasonic sensor to acoustic mine detection will also be addressed.

4:00
IpPAAa10. Doppler ultrasonic detection of targets buried in grass-covered soil. James M. Sabatier (Univ. of Mississippi, Natl. Ctr. for Physical Acoust., 1 Coliseum Dr., University, MS 38677, sabatier@olemiss.edu) and Andi G. Petculescus (Northwestern Univ., Evanston, IL 60208)

An ultrasonic Doppler vibrometer (UDV) is used outdoors to detect vibrating targets buried in grass-covered soil. The sensor head uses two solid-dielectric transducers, in a pitch–catch configuration. A first set of measurements is done using calibrated vibrational sources (shakers), whose vibrational characteristics are known and/or easily predictable. Then, the system is put to the test of detecting a landmine buried in a realistic environment. The target (landmine) is excited either by a mechanical shaker or by a loudspeaker, through acoustic-to-seismic coupling. The wind speed was monitored continuously. Since it is known that wind degrades the UDV signal, efforts were made to perform the experiments in a still environment. The UDV results are compared with those obtained with a laser Doppler vibrometer.

4:15
IpPAAa11. Investigating nonlinearity in acoustic landmine experiments using a clamped-plate soil oscillator. Dang V. Duong and Murray S. Korman (Dept. of Phys., U.S. Naval Acad., Annapolis, MD 21402)

This project investigates the nonlinear effects involved in the detection of plastic landmines using the acoustic-to-seismic coupling techniques developed by Sabatier (linear methods) and Donskoy (nonlinear methods). The soil-plate oscillator is a laboratory apparatus that represents a good physical model for the VS 1.6 and VS 2.2 inert anti-tank plastic landmines. The apparatus consists of a thick-walled cylinder filled with sifted homogeneous soil resting on a thin elastic plate that is clamped to the bottom of the column. Using a loudspeaker (located over the soil) that is driven by a swept sinusoid, tuning curve experiments are performed. The vibration amplitude versus frequency is measured on a swept spectrum analyzer using an accelerometer located on the soil-air interface or under the plate. The “backbone” curve shows a linear decrease in peak frequency versus increasing amplitude. A two-tone test experiment is per-
formed using two loudspeakers generating acoustic frequencies (closely spaced on either side of resonance, typically $-100$ Hz). A collection of combination frequency component profiles (along with the primaries) are measured across the soil surface. In particular, a double peaked profile occurred at $2f_1 + f_2$ corresponding to a common timpani mode. [Work sponsored by U.S. Army Communications-Electronics Command RDEC, NVESD, Fort Belvoir, VA.]

4:30
1pPAa12. False alarms associated with the acoustic-to-seismic detection of buried land mines. Vladimir N. Fokin, Margarita S. Fokina, and James M. Sabatier (Natl. Ctr. for Physical Acoust., 1 Coliseum Dr., University, MS 38655, vfok@olemiss.edu)

Important problems in landmine detection include false alarms and clutter [high values of the acoustic-to-seismic transfer function (TF) in some frequency bands] that mimic the physics of a buried landmine. Many of these high values of the TF are due to the natural variability of the ground. In this paper both the space-frequency variability of the TF($f,x,y$) and the connection with ground variability are discussed. The viscoelastic, layered model of the ground qualitatively explains high values of the TF both in certain spatial regions of the scan site and for some frequencies. This model and the measured spatial dependence of the TF($f,x,y$) were used to model the space-frequency distribution of the acoustic-to-seismic transfer function. Comparison between the calculated and measured transfer functions in the space-frequency volume show that this model can satisfactorily explain variability of the acoustic-to-seismic transfer function. To investigate the spatial variability of ground parameters, the measured TF for a few sites was analyzed. This analysis reveals that different frequency modulation scales exist in the acoustic-to-seismic TF. These different frequency modulation scales are due to spatial dependencies of ground parameters and ground layering. [Work supported by ONR Grant N00014-02-1-0878.]

MONDAY AFTERNOON, 24 MAY 2004

CONFERENCE ROOM E, 1:00 TO 5:30 P.M.

Session 1pPAb

Physical Acoustics and Biomedical Ultrasound/Bioresponse to Vibration:
Sono Et Gravitas: Robert E. Apfel Memorial Session

Lawrence A. Crum, Cochair

Applied Physics Laboratory, University of Washington, 1013 N.E. 40th Street, Seattle, Washington 98105-6698

Ronald A. Roy, Cochair

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

Chair’s Introduction—1:00

Invited Papers

1:10
1pPAb1. Cavitation nucleation. Lawrence A. Crum (Ctr. for Industrial and Medical Ultrasound, Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

For his dissertation research at Harvard, Bob Apfel chose the subject of homogeneous nucleation, and conceived of some ingenious experiments to test existing theories. By selecting a small microdroplet of liquid, he could make the reasonable assumption that no inhomogeneities were present to serve as preferential sites for liquid rupture. However, Bob also studied dirty liquids, as well as very clean ones, and wrote some seminal papers on inhomogeneous nucleation, in which he developed the Golden rule: Know thy liquid! Currently, considerable attention has been devoted to the study of cavitation generation in vivo, particularly in blood, and, for this case, the nucleation conditions are much different than those for normal liquids. In this presentation, I will review some of Bob’s pioneering studies and present some of our latest studies of cavitation inception, both in vitro and in vivo.

1:25
1pPAb2. The oscillation of vapor bubbles. Andrea Prosperetti and Zhizhong Yin (Dept. of Mech. Eng., Johns Hopkins Univ., Baltimore, MD 21218)

Bob Apfel had so many interests that it is impossible—however fitting and desirable—to pay homage to his work as a whole. Some of his early studies were devoted to bubble nucleation at high superheats. In the first part of this paper a recent application of this phenomenon is described. Once a vapor bubble is generated, its subsequent oscillations (free and forced) present analogies and differences with those of a gas bubble: the second part of the paper focuses on this topic. [Work supported by NSF and NASA.]
1:40
1pPAb3. Inspection of the interior of a collapsing bubble via molecular dynamics. Werner Lauterborn, Thomas Kurz, Burkhard Metten, and Daniel Schanz (Drittes Physikalisches Institut, Universität Goettingen, Buergerstr. 42-44, D-37073 Goettingen, Germany, lb@physik3.gwdg.de)

Strongly collapsing bubbles emit shock waves, faint light flashes and induce chemical reactions. At present, the interior of a collapsing bubble is not yet accessible experimentally to reveal the processes behind this behavior. Thus our knowledge has to be advanced theoretically. The standard method relies on solving a set of partial differential equations. However, in the case of small sonoluminescent bubbles, the number of molecules may become too small for a continuum approach. Therefore, the processes within collapsing sonoluminescent bubbles are investigated by molecular dynamics simulations. The temperature, density and pressure distribution within the bubble are calculated. Whereas mass and heat diffusion inside the bubble are automatically accounted for, they must be explicitly introduced across the boundary to the liquid. In particular, also the dissociation of water vapor and chemical reactions of the dissociation products are taken into account. Results are presented for different acoustic driving conditions (sound pressure amplitude and frequency) and different gas compositions including mixtures of noble gases and water vapor. A prediction of the light emission is given to connect the calculations with observable parameters.

1:55
1pPAb4. Quantitative theoretical explanation of Apfel's experimental phase diagrams for sonoluminescing bubbles. Detlef Lohse and Ruediger Toegel (Dept. of Phys., Univ. of Twente, 7500 AE Enschede, The Netherlands)

Robert Apfel had an enormous impact on the research on single bubble sonoluminescence, the light emission of a single sound driven bubble [for a review, see Brenner et al., Rev. Mod. Phys. 74, 425 (2002)]. In 1996, at the ASA Meeting in Hawaii, he posed a challenge to the theoreticians in the field: Make experimentally testable predictions on single bubble sonoluminescence. Apfel collected the predictions and gave a wonderful review talk on the state of the field. Later, he several times came back to that list, comparing the predictions with latest experimental results. Our own predictions those days referred to the phase diagrams of single bubble sonoluminescence. Later Apfel himself, together with Ketterling, measured those phase diagrams experimentally [J. A. Ketterling and R. E. Apfel, Phys. Rev. Lett. 81, 4991 (1998)]. Though qualitatively our 1996 predictions turned out to be correct, a full quantitative model could only be developed recently [R. Toegel and D. Lohse, J. Chem. Phys. 118, 1863–1875 (2003)]. In the presentation we will compare the model predictions with Apfel’s data.

2:10
1pPAb5. From blood to bubbles: Time resolved micro-particle detection and characterization by scattered ultrasound. Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, ronroy@bu.edu)

Robert Apfel believed in the creative application of acoustics technology to difficult problems in biomedical sensing. Much of his work in this area focused on material characterization, with the intention of effecting diagnosis. His early work in blood cell characterization employed acoustic levitation to measure the bulk mechanical properties of human red blood cells. This subsequently paved the way to the use of high-frequency acoustic scattering to yield the compressibility and density of individual blood cells. Technology developed in this later effort was then adapted to the very difficult problem of transient micro-cavitation detection, and the active cavitation detector (ACD) was born. This paper traces this line of work from its origins and, in the process, serves to celebrate Bob Apfel’s peerless ingenuity and irrepressible creativity.

2:25

The mechanical index, MI, resulted from theoretical considerations of the short-pulse acoustic threshold for inertial cavitation in water populated with microbubbles of all sizes [R. E. Apfel and C. K. Holland, Ultrasound Med Biol. 17, 179–185 (1991)]. In this review, the onset of cavitation will be discussed with reference to Robert Apfel’s legacy of theoretical and experimental data. The questions arise: Can the utility of the MI be extended to situations in which the threshold MI is exceeded, thereby allowing for some estimate of the quantification of a potential bioeffect due to microcavitation? Also, can the MI be extended to situations in which pulses are, unlike the original formulation, not short? Is there a theoretical or semi-empirical basis for the MI threshold below which cavitation is unlikely? Can the MI be used to predict gas contrast agent destruction? The possible consequences of gas body activation associated with aerated lung tissue, intestinal gas pockets or encapsulated gas contrast agents represent specific instances of cavitation considerations relevant to clinical practice. Monitoring the real-time display of the MI (mandated by the FDA) helps clinicians evaluate and minimize the potential risks in the use of diagnostic ultrasound instrumentation. [Research supported by National Institutes of Health Grant R29 HL58761.]

2:40
1pPAb7. Active cavitation detection of asymmetrical inertial cavitation. E. Carr Everbach (Dept. of Eng., Swarthmore College, Swarthmore, PA 19081)

The active cavitation detector (ACD) developed in Bob Apfel’s laboratory has often been employed to quantify pressure thresholds for inception of symmetrical inertial cavitation of microbubbles. In the current application, however, a 30-MHz ACD interrogates individual echo-contrast agent bubbles adhering to a Mylar(TM) sheet that are driven into asymmetrical (jet-producing) collapse by a 1-MHz toneburst (>1 MPa pp). The resulting ACD output suggests that asymmetrical bubble collapse is slower than symmetrical
collide, producing less total radiated acoustic power. ACD output mixed with reference sinusoids at 30 MHz and low pass filtered yields Doppler signals that may be useful in quantifying asymmetrical collapses under biomedically relevant conditions, such as on endothelial walls.

2:55

1pPAb8. Microlesions induced by microcavitation during contrast echocardiography.  Douglas Miller, Peng Li, David Gordon, and William Armstrong  (Univ. of Michigan Medical Ctr., Ann Arbor, MI 48109, douglm@umich.edu)

The purpose of this study was to search for histologically identifiable lesions associated with myocardial contrast echocardiography (MCE) in rats. Diagnostic ultrasound scans with 1:4 end-systolic triggering provided a short-axis view of the left ventricle in rats at 1.5 MHz with 1.45-μs pulses of 1.7 Mechanical Index. Two relatively high doses (400 μl/kg) of OptisonTM ultrasound contrast agent were given 5 min apart during 10 min of MCE. One day after scanning, rats were sacrificed and the hearts fixed for histology. Slides were scored blindly by a pathologist, and photomicrographs in the anterior half of the heart sections were characterized by digital image analysis. Microlesions identified by inflammatory infiltrates were scattered primarily over the anterior half of the sections. Pathologically, there was inflammatory cell infiltration in areas of 0.6±0.5% of the sections for shams and 3.6±3.6% for MCE (P<0.01). Analysis of the photographs from the anterior wall found microlesion areas of 0.5±0.8% for shams and 7.4 ± 5.0% for MCE (P<0.02). Diagnostic MCE at high Mechanical Index has a potential for causing microscale lesions in the myocardium by nucleation of microcavitation. [Work supported by NIH Grant EB0338.]

3:10–3:25  Break

3:25


In 1976, research in collaboration with Bob Apfel demonstrated that low-frequency shape oscillations of hydrocarbon drops levitated in water could be driven using modulated radiation pressure. While that response to modulated ultrasound was subsequently extended to a range of systems, the emphasis here is to recall the initial stages of development in Bob Apfel’s laboratory leading to some publications [P. L. Marston and R. E. Apfel, J. Colloid Interface Sci. 68, 280–286 (1979); J. Acoust. Soc. Am. 67, 27–37 (1980)]. The levitation technology used at that time was such that it was helpful to develop a sensitive method for detecting weak oscillations using the interference pattern in laser light scattered by levitated drops. The initial experiments to verify this scattering method used shape oscillations induced by modulated electric fields within the acoustic levitator. Light scattering was subsequently used to detect shape oscillations induced by amplitude modulating a carrier having a high frequency (around 680 kHz) at a resonance of the transducer. Methods were also developed for quantitative measurements of the drop’s response and with improved acoustic coupling drop fission was observed. The connection with research currently supported by NASA will also be noted.

3:40

1pPAb10. Surface-controlled drop oscillations in space.  R. Glynn Holt  (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummings St., Boston, MA 02215)

A series of experiments probing the effects of surfactants was performed by Bob Apfel and his research group in the 1990s. Several laboratory experiments were carried out in uni-axial acoustic levitators. Two experiments were carried out in a triple-axis levitator called the Drop Physics Module, which was carried on Space Shuttle Columbia as part of the First and Second United States Microgravity Laboratory missions. Liquid drops containing aqueous solutions of soluble surfactants were acoustically positioned and deformed (and in some cases rotated) in order to excite shape mode oscillations. The results of these experiments allowed the inference of surface rheological properties (Gibb’s elasticity, surface viscosity coefficients) as functions of surfactant type and concentration. The highlights of this effort will be presented in a semi-technical fashion. [Work supported by NASA.]

3:55

1pPAb11. Apfel’s superheated drop detector.  Francesco d’Errico  (Dept. of Therapeutic Radiol., Yale Univ. School of Medicine, HRT-219, P.O. Box 208040, New Haven, CT 06520-8040)

The introduction of new approaches for radiation dosimetry is rare. A similar breakthrough occurred in 1979, when Robert Apfel invented the superheated drop detector, a miniature relative of the bubble chamber. A fundamental in high-energy particle physics, the bubble chamber utilizes a liquid briefly brought to a transient, radiation-sensitive superheated state by reducing its pressure. Mass boiling of the liquid is prevented by cyclic pressurization, drastically limiting the detection efficiency. In Apfel’s detector, the liquid is kept in a steady superheated state by fractionating it into droplets and dispersing them in an immiscible host fluid, a perfectly smooth and clean container. The approach extends the lifetime of the metastable droplets to the point that practical application in radiation dosimetry is possible. Bubble formation is measured from the volume of vapor or by detecting individual vaporizations acoustically. Various halocarbons are employed and this permits a wide range of applications. Moderately superheated halocarbons are used for neutron measurements, since they are only nucleated by energetic neutron recoil particles. Highly superheated halocarbons nucleate with much smaller energy deposition and are used to detect photons and electrons. This paper reviews the radiation physics of superheated emulsions and their manifold applications.
4:10

1pPAh12. Testing thin film adhesion strength acoustically. Sameer I. Madanshetty (Kansas State Univ., Manhattan, KS 66506, sameer@ksu.edu), Kevin M. Wanklyn, and Hang Ji (Uncopiers Inc., Manhattan, KS 66506)

A new method of measuring the adhesion strength of thin films to their substrates is reported. The method is based on an analogy with the common tensile test of materials. This is an acoustic method that uses acoustic microcavitation to bring about controlled erosion of the thin film. Based on the insonification pressure and the time to complete erosion, the adhesion strength is assessed. The measurements correctly rank order a set of thin film samples of known adhesion strengths.

4:25

1pPAh13. Fabrication of PLGA polymer microspheres for U. S. mediated gene delivery. Rene G. Williamson, William M. Saltzman (Dept. of Biomed. Eng., Yale Univ., 15 Prospect St., Becton 225, New Haven, CT 06511, rgw27@pantheon.yale.edu), and Janet L. Brandsma (Yale Univ., New Haven, CT 06520)

The promises of gene therapy remain unfulfilled because of the lack of a safe and efficient method for transfecting DNA into cells. PLGA has been used as a vehicle for protein, drug, and gene delivery applications because of its biocompatibility and sustained release properties. PLGA polymer microspheres offer advantages of safety and the possibility of sustained intracytoplasmic delivery. The PLGA also protects the plasmid from degradation. Using the double-emulsion microsphere fabrication technique, a new DNA delivery vehicle, comprising of plasmid DNA and octafluoropropane gas encapsulated in PLGA polymer and PVA stabilizer (Sonospheres) was made. The encapsulated gas offers acoustic activity to the microspheres, which enables them to undergo cavitation in an acoustic field. The goal is to lead to increased DNA transfection when these Sonospheres are subjected to an acoustic field in the MHz frequency range. A summary of the fabrication methods and some initial in vitro studies will be presented.

4:40

1pPAh14. Reflector geometry and pressure at the focus of a shock wave lithotripter. Jonathan Iloreta, Prahallad Iyengar, and Andrew Szeri (Dept. of Mech. Eng., Univ. of California, Berkeley, Berkeley, CA 94720-1740, aszeri@me.berkeley.edu)

It has been experimentally shown [Zhong, J. Acoust. Soc. Am. 113, 586–597 (2003)] that refinement of the geometry of the HM-3 lithotripter reflector suppressed the expansion of cavitation bubbles without compromising stone comminution. This effect has been attributed to a change in the rarefaction tail of the pressure wave near the second focus of the original reflector ($F_2$). Following this idea, a numerical model of the reflection and steepening of a pressure wave from an axisymmetric lithotripter has been developed. The model is based on the Euler equations coupled with the Tait equation of state. Preliminary results of the pressure fields produced by numerous reflector shapes are presented. The results show the changes in the rarefaction tail of the pressure wave near $F_2$ for different geometries, thus hinting at the possibility of optimization of the reflector shape for an ideal waveform. [Work supported by NSF.]

4:55


Parameters of ultrasonic aberration can be obtained from power spectra of scattering when individual scattering measurements from which the spectra are estimated have a common aberration and the same nominal geometry. However, the scattering volumes are then confined to a small spatial region and use of finitely many overlapping volumes that result in a nonzero variance is necessary for the measurements. Assuming the scattering is from a spatially uncorrelated medium, the variance of the spectral estimates is expressed as the product of the variance for a single measurement and a reduction factor that depends on the amount of overlap between each volume pair. This factor describes the rate of convergence and the accuracy of the estimates as a function of the number and the overlap of the scattering volumes. Assuming further that the individual volumes are localized by a Gaussian window and that the centers of the volumes are located on orbits of an icosahedral rotation group, the variance is minimized by adjusting the weight and radius of each orbit. Numerical evaluations using orbits formed by icosahedral vertices, face centers, and edge midpoints show that a significant reduction of variance can be achieved from volumes in a confined region.

5:10–5:30
Announcement
Acoustics Research Letters Online (ARLO)
Special Apfel Memorial Edition
Session 1pPP

Psychological and Physiological Acoustics: Compression in Hearing

Sid P. Bacon, Chair
Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona 85287-1908

Chair’s Introduction—2:15
Invited Papers

2:20
1pPP1. Compression in the cochlea. Mario A. Ruggero (Hugh Knowles Ctr., Dept. of Commun. Sci., Northwestern Univ., 2240 Campus Dr., Evanston, IL 60208-3550, mruggero@northwestern.edu)

Cochlear compression consists of the reduction of the amplitude range of the auditory signal, from 6 orders of magnitude at the stapes to (typically) 1–2 orders of magnitude in the spike rate of individual cochlear afferents. Psychophysical discussions of compression generally focus on basilar-membrane (BM) vibrations at the base of the cochlea, which at the characteristic frequency (only) grow with stimulus intensity at remarkably compressive rates but exhibit almost negligible harmonic distortion. In apical cochlear regions, however, BM compression appears to be substantially weaker than at the base, involves significant dc and harmonic distortion, and is not confined to frequencies near the characteristic frequency. Additional stages of compression exist in mechanical-to-electrical transduction in inner hair cells and in the spike generation mechanism. Both the regional variations in the features of BM nonlinearities and the additional compression introduced by proximal stages of signal transformation should be addressed in physiology-based psychophysical models of auditory processing. Ideally, such models should take into account off-frequency listening and should be based on the output of the cochlea, i.e., the activity of the entire array of cochlear afferents, which is better known and more relevant to central auditory processing than BM vibrations. [Work supported by NIH.]

2:45
1pPP2. Behavioral estimates of compression in normal and impaired ears. Christopher J. Plack (Dept. of Psych., Univ. of Essex, Wivenhoe Park, Colchester CO4 3SQ, England, cplack@essex.ac.uk) and Andrew J. Oxenham (MIT, Cambridge, MA 02139-4307)

Over the last 7 years, forward masking techniques have been used to estimate peripheral compression in the human auditory system. The growth of masking and temporal masking curve (TMC) techniques estimate the response to a tone at characteristic frequency (CF) by comparing the masking function (masker level as a function of signal level and masker-signal interval, respectively) for a masker at the signal frequency with the masking function for a masker below the signal frequency (assumed to be processed linearly at high CFs). Compression can also be estimated using the additivity of forward masking technique, in which the effects on signal threshold of combining two equally effective maskers are used to derive the exponent. Overall, the results suggest strong compression in humans (exponent of 0.2) across the range of CFs tested so far (250 to 8000 Hz). The behavioral techniques also suggest that a sensorineural hearing loss of greater than about 50 dB results in an almost complete linearization of the response. However, recent TMC data suggest that less severe hearing losses are associated with a reduction in gain, and a reduction in the level range over which compression is present, but no reduction in the maximum compression.

3:10

Cochlear damage can lead to a reduction in the overall amount of peripheral auditory compression, presumably due to outer hair cell (OHC) loss or dysfunction. The perceptual consequences of functional OHC loss include loudness recruitment and reduced dynamic range, poorer frequency selectivity, and poorer effective temporal resolution. These in turn may lead to a reduced ability to make use of spectral and temporal fluctuations in background noise when listening to a target sound, such as speech. We tested the effect of OHC function on speech reception in hearing-impaired listeners by comparing psychoacoustic measures of cochlear compression and sentence recognition in a variety of noise backgrounds. In line with earlier studies, we found weak (nonsignificant) correlations between the psychoacoustic tasks and speech reception thresholds in quiet or in steady-state noise. However, when spectral and temporal fluctuations were introduced in the masker, speech reception improved to an extent that was well predicted by the psychoacoustic measures. Thus, our initial results suggest a strong relationship between measures of cochlear compression and the ability of listeners to take advantage of spectral and temporal masker fluctuations in recognizing speech. [Work supported by NIH Grants Nos. R01DC03909, T32DC00038, and R01DC00117.]
Hearing aids incorporate multiband compression to compensate for the loudness recruitment that results from sensorineural hearing impairment. No consensus exists in the hearing-aid industry on the best compressor design or on what the design criteria should be. Differences exist in compressor time constants, number of bands, filter shapes, and fitting formula. Such design differences can result in different aided auditory ability as measured by psychoacoustic tests. This research investigated the effect of different compressor designs on fundamental psychoacoustic ability. Multiband compressors with different time constants and different analysis/synthesis filterbanks designs were simulated. Aided performance by the hearing impaired was calculated under each configuration for several psychoacoustic measures, including forward masking, loudness summation, and simultaneous off-frequency masking. Differences in aided performance in these tasks were found: some compressor designs produced aided psychoacoustic ability that resembled normal performance, while other compressor designs did not significantly alter performance from unaided impairment. A psychoacoustic Turing test is proposed that incorporates these aided performance measures as part of a hearing-aid design, a validation strategy, and a comparative technique for evaluating different hearing-aid designs.

4:00–4:15 Break

4:15

1pPP5. Compression, cochlear implants, and psychophysical laws. Fan-Gang Zeng (364 Med. Surge. II, Univ. of California, Irvine, CA 92612)

Cochlear compression contributes significantly to sharp frequency tuning and wide dynamic range in audition. The physiological mechanism underlying the compression has been traced to the outer hair cell function. Electric stimulation of the auditory nerve in cochlear implants bypasses this compression function, serving as a research tool to delineate the peripheral and central contributions to auditory functions. In this talk, I will compare psychophysical performance between acoustic and electric hearing in intensity, frequency, and time processing, and pay particular attention to the data that demonstrate the role of cochlear compression. Examples include both the cochlear-implant listeners’ extremely narrow dynamic range and poor pitch discrimination and their exquisite sensitivity to changes in amplitude and phase. A unified view on the complementary contributions of cochlear compression and central expansion will be developed to account for Webers’ law and Stevens power law.

Contributed Papers

4:40

1pPP6. The compression curves of different models of cochlear nonlinearity. James M. Harte, Stephen J. Elliott (Inst. of Sound and Vib. Res., Univ. of Southampton, Southampton SO17 1BJ, UK), and Henry Rice (Trinity College, Dublin, Ireland)

A widely used method for representing the nonlinear response of the basilar membrane (BM) is a graph of log amplitude of the BM motion against the log amplitude of the sinusoidal driving pressure, known as the input–output level curve. At low sound pressure levels (less than approximately 30 dB) the level of the BM response rises with sound pressure level at a slope of about 1 dB/dB, indicating a linear response. Above this region the slope of the level curve decreases, typically to about 1/2 to 1/3 dB/dB, indicating a compressive nonlinearity. Various models for cochlear nonlinearity will be presented, particularly contrasting the characteristics and behavior of an instantaneously acting nonlinear function and those of level-dependent systems. Both of these models have been used historically to model the nonlinearities present in the cochlea. Comparisons will be made between the properties of these models, highlighting potential methods of distinguishing between them. Similar nonlinear responses to those of the BM are observed in otoacoustic emissions. Experimental results will be reported which attempt to distinguish between the instantaneously acting and level-dependent models of this nonlinearity, which suggest that the latter is more likely.

4:55

1pPP7. The temporal effect in a notched-noise masker for normal-hearing and hearing-impaired listeners. Elizabeth A. Strickland (Dept. of Audiol. and Speech Sci., Purdue Univ., West Lafayette, IN 47907, estrick@purdue.edu)

This is another in a series of studies on the relationship between the temporal effect (or overshoot) and the cochlear active process. The temporal effect refers to the change in threshold for a short-duration tone as it is delayed from the onset of a longer-duration masker. In a previous study, the temporal effect was measured as a function of signal level in a broad-band masker, for listeners with normal hearing and listeners with varying degrees of cochlear hearing loss. When analyzed in terms of cochlear input–output functions, the results were consistent with a decrease in gain at the signal frequency during the course of the masker, as well as a decrease in gain with hearing impairment. In the present study, the temporal effect was measured for a 10-ms, high-frequency signal in a masker with a symmetric notch (relative frequency of 0.3) about the signal frequency, for listeners with normal hearing and cochlear hearing loss. This condition is of interest because the thresholds may reflect the effects of suppression as well as excitatory masking. The relationship between the degree of hearing impairment and the growth of masker level with signal level will be analyzed. [Work supported by the Kinley Trust.]

5:10

1pPP8. Cochlear nonlinearity between 500 and 8000 Hz in listeners with moderate cochlear hearing loss. Enrique A. Lopez-Poveda (Instituto de Neurociencias de Castilla y Leon, Universidad de Salamanca, Avda. Alfonso X El Sabio s/n, Salamanca, Spain, ealopezpoveda@usal.es), Christopher J. Plack, Ray Meddis (Univ. of Essex, Colchester C04 3SQ, UK), and Jose L. Blanco (Oticon Espaa, 28108 Alcobendas, Madrid, Spain)

For two listeners with flat hearing loss, temporal masking curves (TMCs) were measured for probe frequencies (Fp) of 0.5, 1.0, 2.0, 4.0, and 8.0 kHz and masker frequencies of 0.5, 0.6, 0.7, 0.9, 1.0, 1.05, 1.1, and 1.2 Fp. From these, basilar membrane (BM) input/output (IO) functions and psychophysical tuning curves (PTCs) were derived and compared with corresponding data for normal-hearing listeners. Linear IO functions were observed for one ear only and for probe frequencies of 4 kHz and above. This result is consistent with the consequences of severe outer hair cell (OHC) damage on the BM response. Elsewhere, however, IO functions showed residual compression, with slopes in the compression region close to normal. This suggests that either the impairment relates mostly to inner hair cell (IHC) dysfunction, or that OHC dysfunction reduces the gain of the BM at low levels while maintaining the degree of maximum compression. The data suggest that the active and the passive
BM mechanisms have different relative positions (in frequency) in the apical and the basal regions of the cochlea. Across probe frequencies, the slopes of the TMCs for off-frequency maskers are shallower for the impaired ears. It is considered that this possibly reflects IHC-related compression.

5:25

1pPP9. A hearing-aid signal-processing scheme based on the temporal aspects of compression. Laurel H. Carney (Inst. for Sensory Res. and Dept. of Bioeng. and Neurosci., Syracuse Univ., Syracuse, NY, lacarney@syr.edu), Lufeng Shi, and Karen A. Doherty (Syracuse Univ., Syracuse, NY)

Changes in gain associated with the basilar membrane compressive nonlinearity are accompanied by changes in the bandwidth of tuning. Filters with level-dependent bandwidth have level-dependent phase properties. These phase properties result in level-dependent timing of sustained phase-locked responses of auditory-nerve (AN) fibers at low frequencies and level-dependent latencies at high frequencies, where phase-locking rolls off. In the healthy ear, level-dependent temporal aspects of AN responses carry information about stimulus level and spectral properties. Loss of compression with hearing impairment thus results not only in a reduction of amplification, but also in distortion of the temporal response pattern of the AN. The temporal aspects of compression suggest that signal-processing schemes that attempt to correct sounds, or restore normal spatio-temporal response patterns, should include dynamic level-dependent phase shifts. A nonlinear signal-processing scheme will be presented which includes dynamic frequency- and level-dependent phase shifts, based on physiological models of the temporal response properties of AN fibers. Preliminary testing measured listeners preferences for sentences and intelligibility of vowel-consonant syllables with different degrees of nonlinear processing. Hearing-impaired listeners tended to prefer the dynamically corrected stimuli based on improved clarity. Correction also improved intelligibility for some phonemes. [Work supported by NIDCD R21-006057.]

5:40

1pPP10. The effect of amplitude compression, time constants, and number of channels on speech intelligibility in noise. Rolph Houben and Guido F. Smoorenburg (UMC-Utrecht, Heidelberglaan 100, 3508 GA Utrecht, The Netherlands)

The influence of several compression parameters on speech intelligibility in speech-shaped noise was systematically investigated. Experimental conditions include all combinations of compression ratio (CR=1/2, 2/2, 2/3, 3/3 for low/high frequencies), attack and release times (Ta/Tr =4/4, 4/40, 4/400, and 4/400 ms), and number of channels (1, 2, and 6). Twenty subjects with moderate sensorineural hearing loss took part in the experiment. The best average speech reception threshold occurred with two-channel compression at a compression ratio of 2/3 and Ta/Tr =40/40 ms. It was 0.7 dB better than linear amplification. The standard deviation between subjects was 1.7 dB. With six-channel compression the best results (0.4 dB better than linear amplification) were obtained at larger time constants (Ta/Tr=40/400 ms) and CR =2/2 and 2/3. The best result with single-channel compression was equal to linear amplification and occurred for the longest time constant (Ta/Tr=40/400 ms) and CR = 2/2. In fluctuating noise the best speech reception threshold was found with single-channel compression at CR =2/2 and Ta/Tr=40/400 ms. It was 0.9 dB better than linear amplification; the standard deviation was 2.3 dB. The results suggest that finding the optimal condition for an individual is tedious.

5:55–6:30

Panel Discussion

MONDAY AFTERNOON, 24 MAY 2004

IMPERIAL BALLROOM B, 1:00 TO 5:00 P.M.

Session 1pSC

Speech Communication: Poster Session II

Carole E. Gelfer, Chair
Department of Communication Disorders, William Patterson University, Wayne, New Jersey 07470

Contributed Papers

All posters will be on display from 1:00 p.m. 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.

1pSC1. Identification of 20-ms speech samples from normal and neurologically impaired talkers: Do coarticulation differences matter? Joan E. Sussman, Kris Tjaden, and Ya-ju Yu (Dept. of Commun. Disord. and Sci., Univ. at Buffalo, 122 Cary Hall, 3435 Main St., Buffalo, NY 14214)

Prior studies of speech suggest that the extent of coarticulation varies both across and within talkers. Speakers sometimes use average, high, or low amounts of coarticulation. The perceptual consequences of coarticular variability are not well understood, particularly for the speech of individuals with motor control disorders. In the current study, listeners heard the same 20-ms tokens that were used to calculate the degree of anticipatory coarticulation of the following vowel. Acoustic measures were made from the first 20 ms of stop consonants [k] or [t], or from a point 70 ms preceding the vowel in [s] productions. Healthy talkers, speakers with Multiple Sclerosis, and speakers with Parkinson’s disease produced the stimuli. Speech tokens included three degrees of anticipatory coarticulation: average, high, and low. The listeners’ task was to identify whether the vowels that followed the consonant were originally [i] or [u]. Results showed that identification accuracy was poorer for the current 20-ms tokens than for prior speech samples including all aperiodicity of...
consonants. Identification was poorest for [s] and best for [k] tokens. The speed of listener response gave some indication of the differing extents of coarticulation. Perceptual validity for the 20-ms measure of coarticulation is discussed.

IpSC2. Are vowel errors influenced by consonantal context in the speech of persons with aphasia? Carole E. Gelfer (Dept. of Commun. Disord., William Paterson Univ., Wayne, NJ 07470, gelferc@wpunj.edu), Fredericka Bell-Berti (St. John’s Univ., Jamaica, NY 11439), and Mary Boyle (Montclair State Univ., Upper Montclair, NJ 07043)

The literature suggests that vowels and consonants may be affected differently in the speech of persons with conduction aphasia (CA) or non-fluent aphasia with apraxia of speech (AOS). Persons with CA have shown similar error rates across vowels and consonants, while those with AOS have shown more errors for consonants than vowels. These data have been interpreted to suggest that consonants have greater gestural complexity than vowels. However, recent research [M. Boyle et al., Proc. Internat. Cong. Phon. Sci., 3265–3268 (2003)] does not support this interpretation: persons with AOS and CA both had a high proportion of vowel errors, and vowel errors almost always occurred in the context of consonantal errors. To examine the notion that vowels are inherently less complex than consonants and are differentially affected in different types of aphasia, vowel production in different consonantal contexts for speakers with AOS or CA was examined. The target utterances, produced in carrier phrases, were bVC and bV syllables, allowing us to examine whether vowel production is influenced by consonantal context. Listener judgments were obtained for each token, and error productions were grouped according to the intended utterance and error type. Acoustical measurements were made from spectrographic displays.

IpSC3. The effects of complementary and alternative medicine on the speech of patients with depression. Michael Fraas (Dept. of Commun. Sci. and Disord., Univ. of New Hampshire, Durham, NH 03824, michael.fraas@unh.edu) and Michele Solloway (Univ. of New Hampshire, Durham, NH 03824)

It is well documented that patients suffering from depression exhibit articulatory timing deficits and speech that is monotonous and lacking pitch variation. Traditional remediation of depression has left many patients with adverse side effects and ineffective outcomes. Recent studies indicate that many Americans are seeking complementary and alternative forms of medicine to supplement traditional therapy approaches. The current investigation wishes to determine the efficacy of complementary and alternative medicine (CAM) on the remediation of speech deficits associated with depression. Subjects with depression and normal controls will participate in an 8-week treatment session using polarity therapy, a form of CAM. Subjects will be recorded producing a series of spontaneous and narrative speech samples. Acoustic analysis of mean fundamental frequency (F0), variation in F0 (standard deviation of F0), average rate of F0 change, and pause and utterance durations will be conducted. Differences pre- and post-CAM therapy between subjects with depression and normal controls will be discussed.

IpSC4. The effects of Parkinson’s disease on the production of contrastive stress. Henry S. Cheang and Marc D. Pell (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine W., Montreal, QC H3G 1A8, Canada, hchean@po-box.mcgill.ca)

Reduced speech intelligibility has been observed clinically among patients with Parkinson’s disease (PD); one possible contributor to these problems is that motor limitations in PD reduce the ability to mark linguistic contrasts in speech using prosodic cues. This study compared acoustic aspects of the production of contrastive stress (CS) in sentences that were elicited from ten subjects with PD and ten matched control subjects without neurological impairment. Subjects responded to questions that biased them to put emphasis on the first, middle, or last word of target utterances. The mean vowel duration and mean fundamental frequency (F0) of each keyword were then measured, normalized, and analyzed for possible differences in the acoustic cues provided by each group to signal emphatic stress. Both groups demonstrated systematic differences in vowel lengthening between emphasized and unemphasized words across word positions; however, controls were more reliable than PD subjects at modulating the F0 of emphasized words to signal its location in the utterance. Group differences in the F0 measures suggest one possible source of the impoverished intelligibility of Parkinsonian speech and will be investigated in a subsequent study that looks at the direct impact of these changes on emphasis perception by listeners. [Work supported by CIHR.]

IpSC5. Phonological processing among good and poor readers. Ratree Wayland (Prog. in Linguist., Univ. of Florida, 4131 Turlington Hall, Gainesville, FL 32611-5454, ratree@ufl.edu)

Many researchers believe that a connection exists between phonological processing skills and reading ability, and phonological deficits have often been cited as possible explanation for reading disability among both children and adults. This study will present research findings on phonological processing of various speech sounds among school-aged children who were classified as good and poor readers by standardized tests. These subjects will be administered speech discrimination tests using a variety of speech stimuli. Results of their performance on these tasks will be presented and a relationship between their reading and phonological processing abilities will be discussed.

IpSC6. Discrimination and identification of long vowels in children with typical language development and specific language impairment. Hia Datta, Valerie Shafer (Speech and Hearing Sci., The Grad. Ctr., CUNY, 365 Fifth Ave., New York, NY 10016), and Diane Kurtzberg (Albert Einstein College of Medicine, Yeshiva Univ., Bronx, NY)

Researchers have claimed that children with specific language impairment (SLI) have particular difficulties in discriminating and identifying phonetically similar and brief speech sounds (Stark and Heinz, 1966; Studdert-Kennedy and Bradley, 1997; Sassman, 1993). In a recent study (Shafer et al., 2004), children with SLI were reported to have difficulty in processing brief (50 ms), phonetically similar vowels (β/E). The current study investigated perception of long (250 ms), phonetically similar vowels (β/E) in 8- to 10-year-old children with SLI and typical language development (TLD). The purpose was to examine whether phonetic similarity in vowels leads to poorer speech-perception in the SLI group. Behavioral and electrophysiological methods were employed to examine discrimination and identification of a nine-step vowel continuum from /I/ to /E/. Similar performances in discrimination were found for both groups, indicating that lengthening vowel duration indeed improves discrimination of phonetically similar vowels. However, these children with SLI showed poor behavioral identification, demonstrating that phonetic similarity of speech sounds, irrespective of their duration, contribute to the speech perception difficulty observed in SLI population. These findings suggest that the deficit in these children with SLI is at the level of working memory or long term memory representation of speech.

IpSC7. Within-speaker speech intelligibility in dysarthria: Variation across a reading passage. Yana Yunusova, Gary Weismer, John Westbury, and Nicole Rusche (Dept. of Commun. Disord. and Waisman Ctr., Univ. of Wisconsin—Madison, 1500 Highland Ave., Madison, WI 53705)

Understanding factors underlying intelligibility deficits in dysarthria is important for clinical and theoretical reasons. Correlation/regression analyses between intelligibility measures and various speech production measures (e.g., acoustic or phonetic) are often reported in the literature. However, the analyses rarely control for the effect of a third variable
(severity of speech disorder, in this case) likely to be correlated with the primary correlated variables. The current report controls for this effect by using a within-speaker analysis approach. Factors that were hypothesized to underlie the intelligibility variations in multiple breath groups within a connected discourse included structural elements (e.g., number of total words) as well as acoustic measures (e.g., F2 variation). Results showed that speech intelligibility in dysarthric speakers with two forms of neuro- logical disease (Parkinson and ALS) does, in fact, vary across breath groups extracted from a connected discourse, and that these variations are related in some cases to a per breath estimate of F2 variation. [Work supported by NICHD Award No. R01 DC03723.]

IpSC7. Speech proody in Friedrich’s and olivo-ponto cerebellar atrophy. Maureen Casper (Grad. School and Univ. Ctr. of City Univ. of New York, NY)

A critical issue in the study of speech motor control is the identification of the mechanisms that generate the temporal flow of serially ordered articulatory events. Two staged models of serial ordered events (Lashley, 1951; Lindblom, 1963) claim that time controls events whereas dynamic models predict a relative relation between time and space. Each of these models predicts a different relation between the acoustic measures of formant frequency and segmental duration. The most recent method described herein provides a sensitive index of speech deterioration which is both acoustically robust and phonetically systematic. Both acoustic and magnetic resonance imaging measures were used to describe the speech disturbance in two neurologically distinct groups of cerebellar ataxia: Friedreich’s ataxia and olivo-ponto cerebellar ataxia. The speaking task was designed to elicit six different prosodic conditions and four prosodic contrasts. All subjects read the same syllable embedded in a sentence, under six different prosodic conditions. Pair-wise comparisons derived from the six conditions were used to describe (1) final lengthening, (2) phrasal accent, (3) nuclear accent and (4) syllable reduction. An estimate of speech deterioration as determined by individual and normal subjects’ acoustic values of syllable duration, formant and fundamental frequencies was used in correlation analyses with magnetic resonance imaging ratings.


Previously, it was found that 16 right-handed patients with idiopathic Parkinson disease who underwent unilateral implantation of deep brain stimulator in subthalamic nucleus (STN) showed significant improvement in their nonspeech motor functions. Eight of the 16 patients had stimulator in the left STN and eight in the right STN. In contrast, their speech function showed very mild improvement that was limited to the respiratory/phonatory subsystems. Further, there seemed a trend that the patient with right STN stimulation did better than those with left STN stimulation. It was speculated that the difference might be due to a micro lesion caused by the surgical procedure to the corticobulbar fibers run in the left internal capsule. This paper reports speech changes associated with bilateral DBS in STN in four of the 16 subjects who elected to have deep brain stimulator implanted in STN on the opposite side of the brain at a later time. Results show negative changes in speech after bilateral DBS in STN. The changes were not limited to the micro lesion effect due to the surgery itself, but also related to the active stimulation on the dominant hemisphere for speech processing. [Work supported by NIH.]

IpSC10. The effect of noise on the perception of phonetic features in acoustic simulations of cochlear implant speech. Benjamin Munson, Peggy B. Nelson, and Jill E. Mueske (Dept. of Commun. Disord., Univ. of Minnesota, 115 Shevlin Hall, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

Previous research [Munson and Nelson, J. Acoust. Soc. Am. 114, 2360 (2003)] found that adults with cochlear implants (CIs) perceive synthetic /r/-/l/, /w/-/j/, and /s/-/t/ continua less accurately and less categorically than listeners with normal-hearing sensitivity (NH) in noise. Considerable variability was found among the CI listeners: some demonstrated performance comparable to that of NH listeners, while others showed considerably poorer performance. One potential reason for the variability in performance is the number of distinct spectral channels available to the CI listeners. Friesen et al. [J. Acoust. Soc. Am. 110, 1150–1163 (2001)], for example, demonstrated that better-performing implant users showed improvements in consonant and vowel recognition when the number of signal spectral channels was increased, with asymptotic performance at seven channels. Poorer-performing users, however, showed no improvement in scores when the number of channels increased beyond four. To examine this hypothesis, NH listeners participated in an experiment in which they listened to 4- and 8-band acoustic simulations of synthetic continua in quiet and in steady speech-shaped noise at a +10 dB SNR. Analyses will focus on the effect of number of channels on identification accuracy and within-group variability.


The current investigation examined differences between hearing impaired (HI) listeners and normal-hearing listeners (NH) in the amount of masking release (MR) for sentence recognition and syllable identification tasks when they listened in modulated noise. HI and NH listeners’ performance was compared when the performance levels of the two groups were equal in steady noise and in quiet. The relationships between the amount of MR to hearing threshold and suprathreshold abilities of forward masking, auditory filter bandwidth, and auditory stream segregation/integration were also investigated. To compensate for reduced hearing sensitivity for HI listeners, the spectrum levels of both speech and noise were adjusted based on the individual hearing loss. There was no significant performance difference between NH and HI groups in steady noise and in quiet. However, the amount of MR for sentences and for CV syllables was significantly reduced for HI listeners. For sentence recognition, the amount of MR seemed to be more related to hearing sensitivity for low-to-mid frequencies and the characteristics of auditory filters. Performance for gated sentence recognition was also strongly correlated with sentence recognition in gated noise. In contrast, forward masking thresholds appear to be the main contributor to the amount of MR for syllable recognition.

IpSC12. Vowel formant movement and duration perceived through noise vocoders and cochlear implants. Paul Iverson, Bronwen G. Evans, and Charlotte A. Smith (Dept. of Phonet. and Linguist., Univ. College London, 4 Stephenson Way, London NW1 2HE, UK, paul@phonetics.ucl.ac.uk)

Formant movement and duration have been increasingly shown to be important cues for vowel recognition by normal-hearing adults; individuals enhance formant movement and duration contrasts when speaking clearly, and vowel recognition accuracy declines when these differences are reduced in signal-processed or synthesized speech. This study investigated how these cues contribute to vowel recognition by cochlear implant users and normal-hearing individuals listening to noise-vocoded speech. Individuals were tested on vowel recognition using stimuli with and without formant movement and duration cues, and performed a goodness rating task that found best exemplars for vowels that had formant movement and duration variation. The results suggested that duration was an
important cue to recognition for all listeners; removing duration reduced recognition accuracy by 6–19 percentage points in all conditions. Removing formant movement cues reduced the recognition of diphthongs to chance levels, but had little effect on the recognition of monophthongs. Goodness-rating data demonstrated that all listeners preferred systematic patterns of formant movement and duration variation, as occurs in natural speech. The results suggest that cochlear implant users and normal-hearing individuals give similar cue weightings to vowel formant movement and duration, but that formant movement may have less importance under these conditions than in previous studies.

IpSC13. Aided speech recognition in single-talker competition by elderly hearing-impaired listeners. Maureen Coughlin and Larry Humes (Speech and Hearing Dept., Indiana Univ., 200 S. Jordan Ave., Bloomington, IN 47405, mcowghl@indiana.edu)

This study examined the speech-identification performance in one-talker interference conditions that increased in complexity while audibility was ensured over a wide bandwidth (200–4000 Hz). Factorial combinations of three independent variables were used to vary the amount of informational masking. These variables were: (1) competition playback direction (forward or reverse); (2) gender match between target and competition talkers (same or different); and (3) target talker uncertainty (one of three possible talkers from trial to trial). Four groups of listeners, two elderly hearing-impaired groups differing in age (65–74 and 75–84 years) and two young normal-hearing groups, were tested. One of the groups of young normal-hearing listeners was tested under acoustically equivalent test conditions and one was tested under perceptually equivalent test conditions. The effect of each independent variable on speech-identification performance and informational masking was generally consistent with expectations. Group differences in the observed informational masking were most pronounced for the oldest group of hearing-impaired listeners. The eight measures of speech-identification performance were found to be strongly correlated with one another, and individual differences in speech understanding performance among the elderly were found to be associated with age and level of education. [Work supported, in part, by NIA.]

IpSC14. Vowel acquisition by prelingually deaf children with cochlear implants. Marie-Eve Bouchard (Cognit. Neurosci. Ctr., Univ. of Quebec at Montreal, Montreal, QC, Canada), Marie-Thérèse Le Normand (Hôpital de la Salpêtrière, Paris, France), Lucie Ménard, Marilyne Goud, and Henri Cohen (Univ. of Quebec at Montreal, Montreal, QC, Canada)

Phonetic transcriptions (study 1) and acoustic analysis (study 2) were used to clarify the nature and rhythm of vowel acquisition following the cochlear implantation of prelingually deaf children. In the first study, seven children were divided according to their degree of hearing loss (DHL): DHL I: 90–100 dB of hearing loss, 1 children; DHL II: 100–110 dB, 3 children; and DHL III: over 110 dB, 3 children. Spontaneous speech productions were recorded and videotaped 6 and 12 months postsurgery and vowel inventories were obtained by listing all vowels that occurred at least twice in the child’s repertoire at the time of recording. Results showed that degree of hearing loss and age at implantation have a significant impact on vowel acquisition. Indeed, DHL I and II children demonstrated more diversified as well as more typical pattern of acquisition. In the second study, the values of the first and second formants were extracted. The results suggest evolving use of the acoustic space, reflecting the use of auditory feedback to produce the three phonological features exploited to contrast French vowels (height, place of articulation, and rounding). The possible influence of visual feedback before cochlear implant is discussed.

IpSC15. Optimizing cochlear implant frequency boundary tables for vowel perception: A computer simulation. Marias S. Fournakis (Speech & Hearing Sci., Ohio State Univ., 1070 Carmack Rd., Columbus, OH 43210), John W. Hawks (Kent State Univ., Kent, OH 44242, jhawk@kent.edu), and Amy Schwager (Ohio State Univ., Columbus, OH 43210)

For cochlear implants, the assignment of frequency bands to electrodes is a variable parameter that determines what region of the acoustic spectrum will be represented by each electrode’s output. Technology will soon allow for considerable flexibility in programming this parameter. In a first attempt to optimize these assignments for vowel perception, a computer program was written to evaluate different assignment schemes for categorization accuracy based strictly on the frequency values of the first two or three forms, Databases [J. Hillenbrand et al., J. Acoust. Soc. Am. 97, 3099–3111 (1995)] of formant measurements from American English vowels as uttered by men, women, and children were used. For this simulation, it was assumed that each formant frequency was associated with only the frequency band its center frequency fell within. Each pattern of frequency bands was assigned a vowel category identification based on the plurality of tokens whose intended identification category fell within that pattern. A range of frequency scaling schemes for 19 and 20 electrode arrays was evaluated, with the best of these fine tuned for minimum error. The results indicate that manufacturer’s default assignments categorize reasonably well, but a bark-scaled scheme yielded the best unmodified classifications.

IpSC16. An across-frequency deficit in hearing-impaired listeners is supported by acoustic correlation. Eric W. Healy, Anand Kannabiran (Dept. of Commun. Sci. and Disorders., Univ. of South Carolina, Columbia, SC 29208, ewh@sc.edu), and Sid P. Bacon (Arizona State Univ., Tempe, AZ 85287-0102)

It has been recently suggested that listeners having a sensorineural hearing impairment (HI) may possess a deficit in their ability to integrate speech information across different frequencies. When presented with a task that required across-frequency integration of speech patterns and controlled for other known deficits, HI listeners performed more poorly than their normal-hearing (NH) counterparts [C. W. Turner et al., J. Speech Lang. Hear. Res. 42, 773–784 (1999); E. W. Healy and S. P. Bacon, J. Speech Lang. Hear. Res. 45, 1262–1275 (2002)]. The latter study also showed that HI performance fell more steeply when increasing amounts of temporal asynchrony were introduced to the pair of widely separated patterns. In the current study, the correlations between the fluctuating envelopes of the acoustic stimuli were calculated, both when the patterns were aligned and also at various between-band asynchronies. It was found that the rate at which acoustic correlation fell as a function of asynchrony closely matched the rate at which intelligibility fell for the NH listeners. However, the intelligibility scores produced by the HI listeners fell more steeply than the acoustic analysis would suggest. Thus, these data provide additional support for the existence of an across-frequency deficit in HI listeners. [Work supported by NIH.]
normal realm of the AI, which was intended for a much larger mix of sounds, not 16 consonants and a fixed vowel. Most individual consonant \( \text{PF} \) functions obey the same Fletcher AI formula, but with different values of \( e_{\text{min}} \).


Human speech recognition has been studied using response to CV speech stimuli. Miller and Nicely (1955) studied such data in the form of confusion matrices to obtain insight into the psychological structure of the phone in noise. Here, the confusion matrix is modeled as 16 phones, with the transformation matrices as the coordinates in a high dimensional perceptual vector space. The model generalizes to an eigenvalue decomposition (EVD) [Allen (2004)]. This is followed by agglomerative hierarchical clustering of the transformed data, and an automated process is used to identify the main clusters. The resulting EVD clustering is very similar to other Miller–Nicely groupings, based on both production and MDS derived features, but is more model based. It was found that there is a gradual and highly consistent change in the clustering of sounds, independent of cluster size and configuration. By examining the change in similarity between various speech sounds, it is hoped that perceptual features may be uniquely identified.


The accuracy of speech processing applications degrades when operating in co-channel environment. Co-channel speech occurs when more than one person is talking at the same time. The idea of usable speech segmentation is to identify and extract those portions of co-channel speech that are minimally degraded but still useful for speech processing applications (such as speaker identification or speech recognition) which do not work in co-channel environments. Usable speech measures are features that are extracted from the co-channel signal to distinguish between usable and unusable speech. Several usable speech extraction methods have recently been developed based on a single feature of the speech signal being considered. In this paper, however, a new usable speech extraction technique, which sequentially and contextually selects several features of the given signal using the K-nearest neighbor classifier, is being investigated. This new approach considers periodicity and structure based features simultaneously in order to achieve the maximum classification rate, and by observing all the incoming frames, avoids the problem of deciding the amount of data needed to make accurate decisions. A 100% accuracy can be achieved in speech processing applications by using this extracted usable speech segment.

1pSC20. Usable speech processing: A novel approach to processing speech in degraded environments. Brent Y. Smolenski (Temple Univ., 12th and Norris, Philadelphia, PA 19122, bsmolens@temple.edu)

One of the main challenges still plaguing speech processing applications is enabling them to work in operational environments where interference and noise abound. The traditional approach has been to use some form of adaptive filtering operation. However, since the speech is nonstationary, it is possible to extract segments from the speech signal that have a large segmental signal-to-noise ratio (SNR) even when the overall SNR is very low. Such high SNR segments frequently occur during voiced speech, and experiments have shown that, using a speaker identification system, these high SNR segments can be correctly identified even when the speech was not attended to. Hence, the idea of usable speech extraction was to provide the frame size is larger than 30 samples.

It is then shown that a modified sinusoidal model and the Teager energy operator can be combined using a context dependent nonlinear regression technique to obtain a low variance estimate of the segmental SNR values provided the frame size is larger than 30 samples.

1pSC21. Unsupervised learning of broad phonetic classes with a statistical mixture model. Ying Lin (Phonet. Lab., Linguist. Dept., UCLA, Los Angeles, CA 90095-1543, yinglin@ucla.edu)

Unsupervised learning of broad phonetic classes by infants was simulated using a statistical mixture model. A mixture model assumes that data are generated by a certain number of different sources—in this case, broad phonetic classes. With the phonetic labels removed, hand-transcribed segments from the TIMIT database were used in model-based clustering to obtain data-driven classes. Simple hidden Markov models were chosen to be the components of the mixture, with mel-cepstral coefficients as the front end. The mixture model was trained using an expectation-maximization-like algorithm. The EM-like algorithm was initialized by a K-means procedure and then applied to estimate the parameters of the mixture model after iteratively partitioning the clusters. The results of running this algorithm on the TIMIT segments suggested that the partitions may be interpreted as gradient acoustic features, and that to some degree the resulting clusters correspond to knowledge-based phonetic classes. Although such correspondences are rather rough, a careful examination of the clusters showed that the class membership of some sounds is highly dependent on their phonetic contexts. Thus, the clusters may reflect the preliminary phonological categories formed during language learning in early childhood.

1pSC22. The influence of phonological priming on variability in articulation. Molly E. Babel and Benjamin Munson (Dept. of Commun. Disord., Univ. of Minnesota, 115 Shevlin Hall, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

Previous research [Sevald and Dell, Cognition 53, 91–127 (1994)] has found that reiterant sequences of CVC words are produced more quickly when the prime word and target word share VC sequences (i.e., sequences like sit sick) than when they are identical (sequences like sick sick). Even slower production rates are found when primes and targets share a CV sequence (sequences like kick sick). These data have been used to support a model of speech production in which lexical items and their constituent phonemes are activated sequentially. The current experiment investigated whether phonological priming also influences variability in the acoustic characteristics of words. Specifically, we examined whether greater variability in the acoustic characteristics of target words was noted in the CV-related prime context than in the identical-prime context, and whether less variability was noted in the VC-related context. Thirty adult subjects with typical speech, language, and hearing ability produced reiterant two-word sequences that varied in their phonological similarity. The duration, first, and second formant frequencies of the target-words’ vowels were measured. Preliminary analyses indicate that phonological priming does not have a systematic effect on variability in these acoustic parameters.

1pSC23. Noise robust speech recognition with support vector learning algorithms. Hassan H. Namarvar and Theodore W. Berger (Dept. of Biomed. Eng., Univ. of Southern California, OHE-500, Los Angeles, CA 90089-1451, heidarin@usc.edu)

We propose a new noise robust speech recognition system using time-frequency domain analysis and radial basis function (RBF) support vector machines (SVM). Here, we ignore the effects of correlative and nonstationary noise and only focus on continuous additive Gaussian white noise. We then develop an isolated digit/command recognizer and compare its performance to two other systems, in which the SVM is used to replace the multilayer perceptron (MLP) and RBF neural networks. All systems are trained under the low signal-to-noise ratio (SNR) condition.
We obtained the best correct classification rate of 83% and 52% for digit recognition on the TI-46 corpus for the SVM and MLP systems, respectively under the SNR = 0 (dB), while we could not train the RBF network for the same dataset. The newly developed speech recognition system seems to be noise robust for medium size speech recognition problems under continuous, stationary background noise. However, it is still required to test the system under realistic noisy environments to observe whether the system keeps its adaptability and robustness under such conditions. [Work supported in part by grants from DARPA CBS, NASA, and ONR.]

IpSC24. Intelligibility of an ASR-controlled synthetic talking face. 
Catherine Siciliano, Geoff Williams, Andrew Faulkner (Dept. of Phonet. and Linguist., Univ. College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK), and Giampiero Salvi (KTH, 10044 Stockholm, Sweden)

The goal of the SYNFACE project is to develop a multilingual synthetic talking face, driven by an automatic speech recognizer (ASR), to assist hearing-impaired people with telephony communication. Previous multilingual experiments with the synthetic face have shown that time-aligned synthesized visual face movements can enhance speech intelligibility in normal-hearing and hearing-impaired users [C. Siciliano et al., Proc. Int. Cong. Phon. Sci. (2003)]. Similar experiments are in progress to examine whether the synthetic face remains intelligible when driven by ASR output. The recognizer produces phonetic output in real time, in order to drive the synthetic face while maintaining normal dialogue turn-taking. Acoustic modeling was performed with a neural network, while an HMM was used for decoding. The recognizer was trained on the Speech-DAT telephone speech corpus. Preliminary results suggest that the currently achieved recognition performance of around 60% frames correct limits the usefulness of the synthetic face movements. This is particularly true for consonants, where correct place of articulation is especially important for visual intelligibility. Errors in the alignment of phone boundaries representative of those arising in the ASR output were also shown to decrease audio-visual intelligibility. [Work supported by the EU IST Project 2001-33327.]

Gongjun Li and Carol Espy-Wilson (Dept. of Elec. and Computer Eng., Univ. of Maryland at College Park, College Park, MD 20742, gongjun@glue.umd.edu)

In speaker verification, the conventional acoustical models (hidden Markov model and vector quantization) are not able to capture a speaker’s dynamic characteristics. In this paper we describe a novel dynamic acoustical model. The training data are viewed as a concatenation of many speech-pattern samples, and the pattern matching involves a comparison of the pattern samples and the test speech. To reduce the amount of computation, a tree is generated to index the entrance to pattern samples using an expectation and maximization (EM) approach, and leaves in the tree are employed to quantize the feature vectors in the training data. The obtained leaf-number sequences are exploited in pattern matching as a temporal model. We use a DTW scheme and a GMM scheme to match the training data and the test speech. Experimental results on NIST ’98 speaker recognition evaluation data show that the accuracy of speaker verification on 3- and 10-s test speech is raised from 71.1% and 75.2% for a baseline GMM-based system to 80.0% and 82.1% for the dynamic acoustical model, respectively. Furthermore, some pattern samples in the training data are correctly tracked by the test speech.

Janet Slijka (Res. Lab. of Electron., MIT, 36-587, 50 Vassar St., Cambridge, MA 02139)

This research is part of a landmark-based approach to modeling speech perception in which sound segments are assumed to be represented as bundles of binary distinctive features. In this model, probability estimates for feature values are derived from measurements of the acoustics in the vicinity of landmarks. The goal of the current project is to automatically detect the features [high] and [low] for vowel segments based on measurements from average spectra. A long-term and a short-term average spectrum are computed using all vowel regions in the utterance and are used to estimate speaker-specific parameters such as average F0 and average F3 (an indicator of vocal tract length). These parameters are used to estimate F1 using a peak-picking process on the average spectrum at each vowel-landmark. Preliminary results are derived from read connected speech for 738 vowels from 80 utterances (two male speakers, two female speakers). Speaker-independent logistic regression analysis using only average F0 and F1 determines the feature [high] with 73% accuracy and the feature [low] with 84% accuracy. Proposals are made for methods to use additional spectral detail to create a more robust estimate for vowels which show significant formant movement. [Work supported by NIH Grant No. DC02978.]

IpSC27. Spectral variability at the transition between successive phonemes. 
Sorin Dusan (Ctr. for Adv. Information Processing, Rutgers Univ., 96 Frelinghuysen Rd., Piscataway, NJ 08854, sdusan@caip.rutgers.edu)

In an experimental study of identification of truncated Japanese syllables it was stated that a short speech interval (approximately 10 ms) that includes the position of maximum spectral transition between a consonant and a vowel carries the most important information for the perception of the consonant and the syllable [S. Furui, J. Acoust. Soc. Am. 80, 1016–1025 (1986)]. A reduced spectral variability at the transition position could partially explain the increase of information at this position. The current study investigates whether there is a decrease in spectral variability at the transition position between successive phonemes compared with the spectral variability at the phoneme centers. The training part of the TIMIT acoustic-phonetic database, containing sentences in English from 462 American speakers, is used to build 2471 diphone models, based on the 61 symbols used in the database for phonetic transcription. The variability of the mel-frequency cepstral coefficients (MFCCs) is evaluated for various diphone models at the phoneme centers and at the phoneme transition position. Preliminary results suggest that spectral variability is not significantly lower at the phoneme transition positions than that at the phoneme centers in these diphone models.

IpSC28. Automatic speech recognizer based on the Spanish spoken in Valdivia, Chile. 
Maria L. Sanchez (Escuela de Ingenieria Acustica, Facultad de Ciencias de la Ingenieria, Universidad Austral de Chile, General Lagos 2086, Valdivia, Chile), Victor H. Poblete, and Jorge Sommerhoff (Universidad Austral de Chile, Valdivia, Chile)

The performance of an automatic speech recognizer is affected by training process (dependent on or independent of the speaker) and the size of the vocabulary. The language used in this study was the Spanish spoken in the city of Valdivia, Chile. A representative sample of 14 students and six professionals all natives of Valdivia (ten women and ten men) were used to complete the study. The sample ranged in age between 20 and 30 years old. Two systems were programmed based on the classical principles: digitalizing, end point detection, linear prediction coding, cepstral coefficients, dynamic time warping, and a final decision stage with a previous step of training: (i) one dependent speaker (15 words: five colors and ten numbers), (ii) one independent speaker (30 words: ten verbs, ten nouns, and ten adjectives). A simple didactical application, with options to choose colors, numbers and drawings of the verbs, nouns and adjectives, was designed to be used with a personal computer. In both programs, the
tests carried out showed a tendency towards errors in short words with monosyllables like “flor,” and “sol.” The best results were obtained in words with three syllables like “disparar” and “mojado.” [Work supported by Proyecto DID UACH N S-200278.]

IpsC29. Regularized reestimation of stochastic duration models for activity derived from simultaneous EGG data. Analysis of interactions periodic and aperiodic components of the acoustic signal, and (i) voicing, (ii) temporal, spectral and modulation measures from the per-iodic and aperiodic components of the acoustic signal, and (iii) voicing activity derived from simultaneous EGG data. Analysis of interactions observed in British/American English and European Portuguese speech corpora will be compared, and the principal findings discussed.

IpsC30. Measures of voiced frication for automatic classification. Philip J. B. Jackson (Ctr. for Vision, Speech and Signal Processing, Univ. of Surrey, Guildford GU2 7XH, UK), Luis M. T. Jesus (Universidade de Aveiro, 3810-193 Aveiro, Portugal), Christine H. Shadle (Univ. of Southampton, Southampton SO17 1BJ, UK), and Jonathan Pincas (Univ. of Surrey, Guildford GU2 7XH, UK)

As an approach to understanding the characteristics of the acoustic sources in voiced fricatives, it seems apt to draw on knowledge of vowels and voiceless fricatives, which have been relatively well studied. However, the presence of both phonation and frication in these mixed-source sounds offers the possibility of mutual interaction effects, with variations across place of articulation. This paper examines the acoustic and articulatory consequences of these interactions and explores automatic tech-niques for finding parametric and statistical descriptions of these phenomena. A reliable and consistent set of such acoustic cues could be used for phonetic classification or speech recognition. Following work on devoicing of European Portuguese voiced fricatives [Jesus and Shadle, in Mahade et al. (eds.) (Springer-Verlag, Berlin, 2003), pp. 1–8], and the modulating effect of voicing on frication [Jackson and Shadle, J. Acoust. Soc. Am. 108, 1421–1434 (2000)], the present study focuses on three types of information: (i) sequences and durations of acoustic events in VC transitions, (ii) temporal, spectral and modulation measures from the periodic and aperiodic components of the acoustic signal, and (iii) voicing activity derived from simultaneous EGG data. Five random repetitions per sentence were requested for each noise condition. The following acoustic cues were extracted for each word within an utterance: duration, average intensity, peak intensity, average fundamental frequency, and peak fundamental frequency. The ratios of these measures were then compared for semantically salient versus nonsalient words. While we expect to see an overall change in all of these prosodic cues for speech produced in noise (Lombard effect), semantically salient words may exhibit a greater ratio of change in some or all features.

IpsC31. Regularized reestimation of stochastic duration models for phone-classification. Martin J. Russell (Electron., Elec. and Computer Eng., Univ. of Birmingham, Edgbaston, Birmingham B15 2TT, UK) and Philip J. B. Jackson (Univ. of Surrey, Guildford GU2 7XH, UK)

Recent research has compared the performance of various distributions (uniform, boxcar, exponential, gamma, discrete) for modeling segment (state) durations in hidden semi-Markov models used for phone classification on the TIMIT database. These experiments have shown that a gamma distribution is more appropriate than exponential (which is implicit in first-order Markov models), and achieved a 3% relative reduction in phone-classification errors [Jackson, Proc. ICASSP, pp. 1349–1352 (2003)]. The parameters of these duration distributions were estimated once for each model from initial statistics of state occupation (offline), and remained unchanged during subsequent iterations of training. The present work investigates the effect of reestimating the duration models in training (online) with respect to the phone-classification scores. First, tests were conducted on duration models reestimated directly from statistics gathered in the previous iteration of training. It was found that the boxcar and gamma models were unstable, meanwhile the performance of the other models also tended to degrade. Secondary tests, using a scheme of annealed regularization, demonstrated that the losses could be recouped and a further 1% improvement was obtained. The results from this pilot study imply that similar gains in recognition accuracy deserve investigation, along with further optimization of the duration model reestimation proce-dure.

IpsC32. The influence of semantic information on the acoustics of speech in noise. Rupal Patel, Mariam Syeda, and Aviva Krauthammer (Dept. of Speech Lang. Pathol. and Audiol., Northeastern Univ., 360 Huntington Ave., 102 FR, Boston, MA 02115, r.patel@neu.edu)

While there is a significant body of work on how people modify their speech patterns in the presence of noise, the role of semantic information on these acoustic modifications is not well understood. This study exam-ined whether adult speakers of English differentially modify semantically salient versus nonsalient words within a sentence in the presence of noise. Participants were asked to produce a set of 20 sentences [from the Speech Perception in Noise List, Kalikow et al. (1977)], in each of five noise conditions: quiet, 60 and 90 dB SPL multispeaker conversation, and 60 and 90 dB SPL street noise. Five random repetitions per sentence were requested for each noise condition. The following acoustic cues were extracted for each word within an utterance: duration, average intensity, peak intensity, average fundamental frequency, and peak fundamental frequency. The ratios of these measures were then compared for semantically salient versus nonsalient words. While we expect to see an overall change in all of these prosodic cues for speech produced in noise (Lombard effect), semantically salient words may exhibit a greater ratio of change in some or all features.

IpsC33. The relation between semantics and lexical properties in spoken word production. Patricia Amico, Jan Charles-Luce, and Elizabeth McEldowney (Dept. of Commun. Disord. and Sci., 122 Cary Hall, Univ. at Buffalo, Buffalo, NY 14214)

Previous research has demonstrated facilitation in speech production when multiple levels are activated. In the present study a time course of facilitation and the contribution of various lexical and semantic properties in facilitating spoken word production was investigated. An experiment was conducted that examined the effects of semantics and lexical properties on acoustic-phonetic duration on spoken word production. Specifi-cally, the primary interest was how duration changed as a function of semantic context and its interaction with the frequency and similarity neighborhood of CVC words. The semantic contexts and targets were presented using a visual naming paradigm. Target words were presented either 100 or 1500 ms after the offset of a simultaneous presentation of a string of three primes. The three primes were either all semantically related to the target, all semantically unrelated to the target, or all nonlin-guistic characters neutral to the target. Duration of the target stimuli and reaction times to onset of articulation were measured. The results will be discussed in terms of interactive activation. [Work supported by NIH NIDCD Grant R01 0265801.]
IpSC34. Adaptation to structural modifications of the human vocal tract during speech: Electropalatographic measures. Wendi A. Aasland, Shari R. Baum (School of Commun. Sci. and Disorder, McGill Univ., 1266 Pine Ave. W., Montreal, QC H3G 1A8, Canada, wendi.aasland@mail.mcgill.ca), and David H. McFarland (Université de Montréal, Montréal, QC, Canada)

Structural modifications to the vocal tract force speakers to alter their previously learned articulatory patterns in order to produce perceptually adequate speech. Previous research has shown that acoustic output in the production of alveolar consonants changes during adaptation to structural alterations of the palate, but to date, little is known regarding exactly how these changes result kinematically. The present study examines the adjustments made to tongue–palate contact patterns, measured using electropalatography (EPG), during adaptation to a palatal perturbation for the fricative [s]. Productions of the nonsense word [asa] were elicited in nine subjects at five time intervals, 15 min apart, while speakers wore electropalates modified with a thicker-than-normal alveolar ridge. Between measurement intervals, speakers read [s]-laden passages to promote adaptation. Productions were also elicited with an unperturbed electropalate in place to characterize normal articulation. Electropalatographic analyses revealed a posterior shift in center of gravity of tongue–palate contact, alterations in the width of the medial groove necessary for [s] production, and increased variability in productions, which may reflect the instability of the new motor programs. Results are discussed in relation to the development of adaptive articulatory programs in speech motor control. [Work supported by NSERC and a FRQSI Bourse de Formation.]

IpSC35. TADA: An enhanced, portable Task Dynamics model in MATLAB. Hosung Nam, Louis Goldstein (Haskins Labs. & Yale Univ., 270 Crown St., New Haven, CT 06511, hosung.nam@yale.edu), Elliot Saltzman (Boston Univ., Boston, MA 02215), and Dani Byrd (USC Linguist. and Haskins Labs., Los Angeles, CA 90089-1693)

A portable computational system called TADA was developed for the Task Dynamic model of speech motor control [Saltzman and Munhall, Ecol. Psychol. 1, 333–382 (1989)]. The model maps from a set of linguistic gestures, specified as activation functions with corresponding constriction goal parameters, to time functions for a set of model articulators. The original Task Dynamic code was ported to the (relatively) platform-independent MATLAB environment and includes a MATLAB version of the Haskins articulatory synthesizer, so that articulator motions computed by the Task Dynamic model can be used to generate sound. Gestural scores can now be edited graphically and the effects of gestural score changes on the models output evaluated. Other new features of the system include: (1) A graphical user interface that displays the input gestural scores, output time functions of constriction goal variables and articulators, and an animation of the resulting vocal-tract motion; (2) Integration of the Task Dynamic model with the prosodic clock-slowing, p-i-gesture model of Byrd and Saltzman [J. Phonetics 31, 149–180 (2003)]. This now allows prosody-driven slowing to be applied to the full set of active gestures and its effects to be evaluated perceptually. [Work supported by NIH.]

IpSC36. The role of coda consonants in triggering speech errors: An ultrasound study. Marianne Pouplier (Vocal Tract Visualization Lab, UMAB Dental School, BMS, 666 W. Baltimore St., Baltimore, MD 21201, mpoupl01@umaryland.edu)

Recent speech-error research using articulatory data has shown that errors can often result in a phonologically illegal structure. These findings have been interpreted to mean that speech production is a fundamentally coordinative process. In speech errors, gestures can fall into grammatically illegal, albeit dynamically stable coordination modes. Shared gestural structures, such as identical consonants in the vicinity of the gestures affected by error, set up the conditions under which dynamically stable coordination modes can come to dominate over grammatical coordination patterns (e.g., Pouplier, 2003). This approach finds support in the often-observed phenomenon that a shared final consonant between two words will increase the likelihood of the initial consonants interacting. The present study uses ultrasound data of speech errors to investigate in what domains other than the coda shared gestural structure can trigger errors. Stimuli without coda consonants are employed to investigate whether errors on the initial consonants can be triggered by a shared vowel only. It is further examined whether the shared structure triggering errors can reside within the initial consonants themselves, e.g., when the initial consonants are complex, multigestural constellations that overlap in some aspects of their gestural composition. [Work supported by NIH R01-DC01758.]

IpSC37. Motor learning of volitional nonspeech oral movements: Intraoral pressure and articulatory kinematics. Susan Shaiman, Malcolm R. McNeil (Dept. of Commun. Sci. & Disorder, Univ. of Pittsburgh, 4033 Forbes Tower, Pittsburgh, PA 15260, shaiman@csd.pitt.edu), and Neil J. Szuminski (Necessity Consulting, Mars, PA 16046)

This study provides initial validation of a complex nonspeech task that will be used in future research for examining and comparing the neurophysiologic mechanisms of speech and volitional nonspeech oral movements. Motor learning of a complex sequence of lip, jaw, and tongue movements was explored as speakers produced one of three intraoral air-pressure targets during each bilateral closing gesture in the sequence. Motor learning was demonstrated by retention of pressure targets subsequent to acquisition. Transfer to other nonspeech tasks was also explored, as were changes in articulatory kinematics with learning. The current non-speech task was constructed to parallel speech production, by controlling several commonly observed physiologic characteristics of speech (i.e., complex sequence of potentially overlapping articulatory segments; goal of intraoral pressure during bilabial closure). This construction of the nonspeech task, along with its demonstrated motor learning, greatly extends the ability to make valid comparisons between speech and nonspeech productions. [Work supported by CMRF-University of Pittsburgh.]

IpSC38. Sensorimotor adaptation to acoustic perturbations in vowel formants. Virgilio Villacorta, Joseph Perkell (Res. Lab. of Electron., MIT, 50 Vassar St., Cambridge, MA 02139, virgilio@mit.edu), and Frank Guenther (Boston Univ., Boston, MA 02215)

The goal of this research is to study the auditory component of feedback control in speech production. This experiment investigates auditory sensorimotor adaptation (SA) as it relates to speech production: the process by which speakers alter their speech production in order to compensate for perturbations of their normal auditory feedback. Specifically, the first formant frequency (F1) was shifted in the auditory feedback heard by naive adult subjects as they produced vowels in single syllable words. Initial results indicate that subjects demonstrate compensatory formant shifts in their speech. This compensation was also present after training when acoustic feedback was masked by noise. This suggests that internal models used in the control of speech movements can be constantly updated by auditory feedback. These results in voiced speech are consistent with results from Houde and Jordan [Science 279, 1213–1216 (1998)], which demonstrated SA in whispered speech. A second study, currently underway, investigates perceptual discrimination of vowel stimuli differing in F1 frequency, using the same subjects as in the SA studies. Cross-subject relations between discrimination scores and extent of compensation will be presented and discussed. [Work supported by NIDCD Grant R01-DC01925.]


Statistical analysis of data variability in speech production research has traditionally been addressed with the assumption of normally distributed error terms. The correct and valid application of statistical procedure requires a thorough investigation of the assumptions that underlie the
methodology. In previous work [Kollia and Jorgenson, J. Acoust. Soc. Am. 102 (1997); 109 (2002)], it was shown that the error terms of speech production data in a linear regression can be modeled accurately using a quadratic probability distribution, rather than a normal distribution as is frequently assumed. The measurement used in the earlier Kollia–Jorgenson work involved the classical Kolmogorov–Smirnov statistical test. In the present work, the authors further explore the problem of analyzing the error terms coming from linear regression using a variety of known statistical tests, including, but not limited to chi-square, Kolmogorov–Smirnov, Anderson–Darling, Cramer–von Mises, skewness and kurtosis, and Durbin. Our study complements a similar study by Shapiro, Wilk, and Chen [J. Am. Stat. Assoc. (1968)]. [Partial support provided by PSC-CUNY and NSF to Jay Jorgenson.]

IpSC40. Spectral characteristics of speech with fixed jaw displacements. Nancy P. Solomon, Matthew J. Makashay (Army Audiol. and Speech Ctr., Walter Reed Army Med. Ctr., 6900 Georgia Ave. NW, Washington, DC 20307, nancy.solomon@na.amedd.army.mil), and Benjamin Munson (Univ. of Minnesota, Minneapolis, MN 55455)

During speech, movements of the mandible and the tongue are interdependent. For some research purposes, the mandible may be constrained to ensure independent tongue motion. To examine specific characteristics of speech with different jaw positions, ten normal adults produced sentences with multiple instances of /t/, /s/, /f/, /l/, /ai/, and /bd/. Talkers produced stimuli with the jaw free to vary, and while gently biting on 2- and 5-mm bite blocks unilaterally. Spectral moments of /s/ and /f/ frication and /t/ bursts differed such that mean spectral energy decreased, and diffuseness and skewness increased with bite blocks. The specific size of the bite block had minimal effect on these results, which were most consistent for /s/. Formant analysis for the vocoids revealed lower F2 frequency in /s/ and at the end of the transition in /ai/ when bite blocks were used; F2 slope for diphthongs was not sensitive to differences in jaw position. Two potential explanations for these results involve the physical presence of the bite blocks in the lateral oral cavity, and the oromotor system’s ability to compensate for fixed jaw displacements. [Work supported by NIDCD R03-DC06096.]

IpSC41. Segment sequencing and the jaw cycle. Melissa Redford (Dept. of Linguist., 1290 Univ. of Oregon, Eugene, OR 97403, redford@darkwing.uoregon.edu) and Paul van Donkelaar (Univ. of Oregon, Eugene, OR)

The open–close jaw cycle established in running speech often spans the articulation of several segments. Moreover, the extremes of the cycle are associated with opposing segment types, consonants and vowels, which form higher-level linguistic units, syllables, when grouped. The present study sought to determine how the syllable is related to the cycle. A partial relationship may imply that the jaw moves linearly from one segment to another. An isomorphic relationship that jaw movement is produced over the articulation of a group of segments, regardless of segment sequencing within the group. Five speakers repeated one of five syllable types (CV, VC, CVC, CCV, VCC) continuously for several seconds, producing an unbroken sequence of consonants and vowels. Jaw movement and acoustic data were recorded simultaneously and used to determine the temporal boundaries of the cycles and syllables, respectively. Cycle boundaries were defined by the points of maximum closure preceding and following an opening gesture, and syllable boundaries according to changes in frequency and amplitude corresponding to the onset of the initial and final segments in the syllable. A comparison of the relative boundary locations suggested an isomorphic relationship between cycles and syllables that interacted with segment type, syllable shape, and speaker.

IpSC42. A new 3D dynamical biomechanical tongue model. Jean-Michel Gerard, Pascal Perrier (ICP-INPG, 46 av. Felix Viallet, 38031 Grenoble, Cedex 1, France, gerard@icp.inpg.fr), Yohan Payan (Univ. Joseph Fourier, 38706 La Tronche, Cedex, France), and Reiner Wilhelms-Tricarico (MIT, Cambridge, MA)

A new dynamical biomechanical tongue model is being developed to study speech motor control. In spite of its computational complexity, a 3D representation was chosen in order to account for various contacts between tongue and external structures such as teeth, palate, and vocal tract walls. A fair representation of tongue muscle anatomy is provided, by designing the finite element mesh from the visible human data set (female subject). Model geometry was then matched to a human speaker, so that simulations can be quantitatively compared to experimental MRI data. A set of 11 muscles is modeled, whose role in speech gestures is well established. Each muscle is defined by a set of elements whose elastic properties correspond to muscle activation. Muscles forces are applied to the tongue model via macrotubes defined within the mesh by muscle specific sets of nodes. These forces are currently specified as step functions. Boundary conditions are set using zero-displacement nodes simulating attachments of tongue on bony structures. The nonlinear mechanical properties of tongue soft tissues are modeled using a hyperelastic material. Three-dimensional tongue deformations generated by each muscle, using FEM software ANSYS for computation, will be presented. Implications for speech motor control will be proposed.


We present novel results from the acoustic and articulatory investigation of the production of the transparent vowels (TVs) /i/, /iː/, /eː/ in Hungarian (colon denotes length). The acoustic measurements of the front–back distinction (second formant, the difference of the first and second formants [Ladefoged, 1993]) show that the effect of adjacent back vowels on the front quality of the TVs is only weakly significant. The articulatory measurements of the same data, however, show that adjacent back vowels cause highly significant retraction of the tongue body during the production of the front TVs. The significance of this finding lies in its relevance to the relationship between phonetics and phonology. Our results demonstrate that minor phonetic differences in articulation, impossible to access by traditional theory, correlate with full-fledged phonological alternation of suffix selection in Hungarian. Traditional phonological accounts predict no effect of continuous phonetic details on discrete phonological generalizations. This is supported in our acoustic data but contrasts with our articulatory findings. In the paper we propose a dynamic model where phonological transparency is directly related to nonlinearity between acoustics and articulation [Stevens, 1989; Wood, 1979]. [Work supported by NIH.]
**Session 1pSP**

**Signal Processing in Acoustics: Distinguished Lecture on Communication Acoustics**

Ning Xiang, Chair  
*Architecture, Rensselaer Polytechnic Institute, 110 8th Street, Troy, New York 12180*

**Chair’s Introduction—1:15**

**Invited Paper**

1:20

1pSP1. **Communication acoustics.**  Jens Blauert  (Inst. of Commun. Acoust., Ruhr-Universitaet Bochum, D-44780 Bochum, Germany, jens.blauert@rub.de)

Those aspects of acoustics which concern the relations of acoustics to the information and communication technologies are now frequently called “communication acoustics.” After a short review of the history of this field, relevant results from recent research at the Institute of Communication Acoustics at the Ruhr-University of Bochum, Germany, will be reported. This work can be seen in light of the research areas of computational auditory scene analysis (CASA) and auditory virtual environments (AVE)—both dealing with the parametric representation of auditory scenes. Recent application opportunities and future trends will be discussed. It will be argued that modern communication acoustical systems—which are often only embedded components in more complex communication systems—require more and more built-in explicit knowledge. Among other things, the development of such components and systems calls for data and knowledge from the cognitive sciences. Today’s programs for education in communication engineering and communication acoustics often do not take sufficient account of this trend.

---

**Session 1pST**

**ASA Committee on Standards: Role of Standards in ASA**

Paul D. Schomer, Chair  
*Scomer & Associates, Inc., 2117 Robert Drive, Champaign, Illinois 61821*

**Chair’s Introduction—1:15**

**Invited Papers**

1:20

1pST1. **Acoustical standards in the Society: Historical perspective.**  Tony F. W. Embleton  (80 Sheardown Dr., Nobleton, ON L0G 1N0, Canada), Paul D. Schomer  (Schomer and Assoc., Inc., Champaign, IL 61821), and Susan B. Blaeser  (Acoust. Society of America Standards Secretariat, Melville, NY 11747)

The Acoustical Society of America has been active in national standards since 1932 when it asked the American Standards Association to initiate a project to standardize acoustical measurements and terminology. The new committee was assigned to the Society and designated Z24. By 1942, Z24 had expanded to include vibration. By 1957 the activity had grown so much that Z24 was split into three committees: S1 Acoustics, S2 Mechanical Vibration and Shock, and S3 Bioacoustics. (S12 Noise was added later, in 1981). Until 1969, these committees were administered by the USA Standards Institute, the predecessor of the American National Standards Institute (ANSI), and new standards were approved by the Executive Council of the Society. In 1971, the Society assumed the responsibility for developing the US position on international standards in acoustics being generated by TC108 of the International Organization for Standardization. During the early 1970s ANSI became the body to approve all national standards, and the Society became a standards developer. ASA established a Committee on Standards (ASACOS) in 1978 to guide the work of the S-Committees and the Standards Secretariat. In the past few years ASACOS has increasingly improved its links with the Society’s other technical activities.
1pST2. Standards 101: The ASA Standards program. Paul Schomer (Schomer and Assoc., Inc., 2117 Robert Dr., Champaign, IL 61821)

ASA serves as a standards developer under the auspices of the American National Standards Institute (ANSI). The Standards Program is organized through four technical committees (S1, S2, S3, and S12) and one administrative committee (ASACOS). S1 deals with physical acoustics, S2 deals with shock and vibration, S3 deals with physiological and psychological acoustics and S12 deals with noise. ASACOS is the ASA Committee on Standards. The program has three primary tasks: (1) development of national standards (ANSI Standards), (2) national adoption of international standards (ANSI NAIS Standards), (3) providing the USA input to the development of international standards (ISO and IEC Standards). At every level the main work is accomplished in Working Groups (WG) that are staffed by hundreds of volunteers, mainly ASA members from its various technical committees such as Noise, Physical Acoustics, Architectural Acoustics, Physiological and Psychological Acoustics, etc. Overall, the Standards Program involves more ASA members than does any other single function of the society except meetings. It is the biggest outreach function of ASA affecting the health, welfare, and economic well-being of large sectors of society. It is a main way the ASA diffuses the knowledge of acoustics and its practical application, perhaps the main way.

2:00

1pST3. The roles of major international standards organizations. Steven Cornish (Director of Intl. Policy, ANSI, 25 W. 43rd St., New York, NY 10036)

This review paper describes and analyzes the roles of the major international standards organizations, the International Organization for Standardization (ISO) and the International Electro-technical Commission (IEC). In particular, it analyzes the relationships of ISO and IEC with ANSI, CEN (including the ISO/CEN Vienna Agreement), and CENELEC. This paper also addresses recent developments to ensure the global relevance of ISO and IEC standards. Lastly, this paper addresses the myths, perceptions, and realities of European participation and potential dominance in ISO.

2:20

1pST4. The role of standards in noise. Richard J. Peppin (Scantek, Inc., 7060 Oakland Mills Rd. #L, Columbia, MD 21046, peppinin@asme.org)

This paper discusses how the noise community uses standards. The noise community consists of people working in the control and abatement of noise sources: produced by, or isolated as a result of, mechanical, information, transportation, vehicle, highway, industrial, and building systems. For the purposes of this paper, standards are described as “agreed upon ways of doing things.” Many of the standards available, their organizations, and their interrelationships will be addressed. The presentation will show how the various disciplines are concerned with the different groupings of standards and why, in many instances, it is impossible to do work without them.

2:40

1pST5. Role of standards in ASA: Animal bioacoustics. Ann E. Bowles (Hubbs-SeaWorld Res. Inst., 2595 Ingraham St., San Diego, CA 92109, annb1@san.rr.com)

Animal bioacoustics (AB) is a participant in ASACOS committee S3, Bioacoustics, and has one working group S3/WG90. However, standards that could be written for animals cut across committee boundaries, from development of terminology (S1), to specification of audiometric methods (S2), to hearing conservation programs (S12). From a biologist’s perspective, there will be a fundamental difference between developing AB standards and those that have been published during the 75-year history of ASA—instead of focusing on one species, Homo sapiens, AB standards will be applied to many, even if data are available for only a few. Given the intensive research dedicated to establishing the existing standards for H. sapiens, the task ahead appears daunting. How should standards for thousands of species be specified? How should the standards process include a measure critical for wild animals, population sustainability? Writers of AB standards must apply a powerful conceptual tool, the comparative approach, and must design standards to incorporate new data quickly. Funding sources must recognize the need for these approaches. If they do, the standards developed will be marketable; it is also likely that important new perspectives on human bioacoustics will emerge.

3:00–3:10 Break

3:10

1pST6. The application of standard definitions of sound to the fields of underwater acoustics and acoustical oceanography. William M. Carey (College of Eng., Boston Univ., Boston, MA 02215, kerrygrp@ctol.net)

Recent societal concerns have focused attention on the use of sound as a probe to investigate the oceans and its use in naval sonar applications. The concern is the impact the use of sound may have on marine mammals and fishes. The focus has changed the fields of acoustical oceanography (AO) and underwater acoustics (UW) because of the requirement to communicate between disciplines. Multiple National Research Council publications, Dept. of Navy reports, and several monographs have been written on this subject, and each reveals the importance as well as the misapplication of ASA standards. The ANSI-ASA standards are comprehensive, however not widely applied. The clear definition of standards and recommendations of their use is needed for both scientists and government agencies. Traditionally the U.S. Navy has been responsible for UW standards and calibration; the ANSI-ASA standards...
have been essential. However, recent changes in the Navy and its laboratory structure may necessitate a more formal recognition of ANSI-ASA standards and perhaps incorporation of UW-AO in the Bureau of Standards. A separate standard for acoustical terminology, reference levels, and notation used in the UW-AO is required. Since the problem is global, a standard should be compatible and cross referenced with the International Standard (CEI/IEC 27-3).

3:30

Experience gained during the external accreditation of the Acoustical Standards Program at the Institute for National Measurement Standards of the National Research Council is discussed. Some highlights include the preparation of documents for calibration procedures, control documents with attention to reducing future paper work and the need to maintain documentation or paper trails to satisfy the external assessors. General recommendations will be given for laboratories that are contemplating an external audit in accordance to the requirements of ISO/IEC 17025.

3:50

The Engineering Acoustics Technical Committee is concerned with the evolution and improvement of acoustical techniques and apparatus, and with the promotion of new applications of acoustics. As cited in the Membership Directory and Handbook (2002), the interest areas include transducers and arrays; underwater acoustic systems; acoustical instrumentation and monitoring; applied sonics, promotion of useful effects, information gathering and transmission; audio engineering; acoustic holography and acoustic imaging; acoustic signal processing (equipment and techniques); and ultrasound and infrasound. Evident connections between engineering and standards are needs for calibration, consistent terminology, uniform presentation of data, reference levels, or design targets for product development. Thus for the acoustical engineer standards are both a tool for practices, for communication, and for comparison of his efforts with those of others. Development of many standards depends on knowledge of the way products are put together for the market place and acoustical engineers provide important input to the development of standards. Acoustical engineers and members of the Engineering Acoustics arm of the Society both benefit from and contribute to the Acoustical Standards of the Acoustical Society.

MONDAY EVENING, 24 MAY 2004

NEW YORK BALLROOM A, 7:00 TO 9:00 P.M.

Session 1eID

Interdisciplinary: Tutorial Lecture: Listening to the Acoustics in Concert Halls

Patricia K. Kuhl, Chair
University of Washington, UW Box 357920, Seattle, Washington 98195

Chair’s Introduction—7:00

Invited Paper

7:05
1eID1. Listening to the acoustics in concert halls. Leo L. Beranek (975 Memorial Dr., Ste. 804, Cambridge, MA 02138-5755) and David Griesinger (Lexicon, Bedford, MA 01730-1441)

How does acoustics affect the symphonic music performed in a concert hall? The lecture begins with an illustrated discussion of the architectural features that influence the acoustics. Boston Symphony Hall, which was built in 1900 when only one facet of architectural design was known, now rates as one of the world’s great halls. How this occurred will be presented. Music is composed with some acoustical environment in mind and this varies with time from the Baroque to the Romantic to the Modern musical period. Conductors vary their interpretation according to the hall they are in. Well-traveled listeners and music critics have favorite halls. The lecture then presents a list of 58 halls rank ordered according to their acoustical quality based on interviews of music critics and conductors. Modern acoustical measurements made in these halls are compared with their rankings. Music recordings will be presented that demonstrate how halls sound that have different measured acoustical parameters. Photographs of a number of recently built halls are shown as examples of how these known acoustical factors have been incorporated into architectural design.