DTU Library

Early reflection energy in concert halls: how much, how early, and from where (A)

Gade, Anders Christian

Link back to DTU Orbit

PROGRAM OF

The 141st Meeting of the Acoustical Society of America

Palmer House Hilton Hotel • Chicago, Illinois • 4–8 June 2001

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, 4 JUNE 2001   CRYSTAL ROOM, 8:30 TO 11:30 A.M.

Session 1aAA

Architectural Acoustics: Concert Hall Measurements

J. Christopher Jaffe, Chair

Jaffe Holden Acoustics, 114A Washington Street, Norwalk, Connecticut 06854

Chair's Introduction—8:30

Invited Papers

8:35

1aAA1. Measurement of decay times in coupled spaces. Ning Xiang (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677), Paul M. Goggans, and Donghua Li (Univ. of Mississippi, University, MS 38677)

To achieve enough variability for different performances, a number of new concert halls with secondary hard chambers coupled to the main floor have recently received positive evaluations. As a result, the acoustics of coupled spaces is drawing growing attention in architectural acoustics. In these coupled spaces, the sound energy decays with different exponential constants under certain conditions. For better understanding and control of the acoustics in coupled spaces an efficient measurement technique of identifying multiple decay times is therefore of critical importance. A Bayesian parameter estimation approach is applied to an algorithmic evaluation of decay times using measured Schroeders decay functions. This work discusses Bayesian analysis of systematic measurements in scaled-down, coupled rooms and real coupled spaces.

8:55

1aAA2. Stage acoustics as affected by an array of apertures incorporated into the orchestra shell of a stagehouse coupled auditorium. Jason E. Summers and Yasushi Shimizu (Sonics in Architecture, Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180, summej2@rpi.edu)

A number of performance halls incorporate into their design a volume (or ''chamber'') surrounding the stage platform and coupled to the stage platform through the orchestra shell. Bass Performance Hall utilizes this design concept, while incorporating into the orchestra shell an array of apertures backed by adjustable doors. Measurements have been made to assess the impact of the adjustability of these doors on the soundfield of the stage platform. Consideration was given to both the early and late portions of the soundfield. Variations in impulse response derived measures, in the spectra of early reflections, and in the coupling behavior are presented. These are examined in view of existing models for the reflection and transmission properties of apertures and for the decay of coupled spaces. [Research supported by the Bass Foundation.]

9:15

1aAA3. Concert hall acoustics at Bass Hall. Derrick Knight, Young-Min Kwon, John Kreiger, Brandon Holcomb, Jamin Hemenway, Jason Summers, Yasushi Shimizu, and Christopher Jaffe (Sonics in Architecture, Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180, knight427@yahoo.com)

Traditionally, multi-purpose halls are believed to have less than adequate symphonic acoustic listening characteristics compared to one room concert halls such as Boston Symphony Hall. Over the last two years, students and faculty of the Sonics Program at the Rensselaer Polytechnic Institute conducted a series of measurements at the multi-purpose Bass Performance Hall in Forth Worth, Texas. In this investigation, Bass Hall is shown to have excellent concert hall metrics when configured for that purpose. Measurements

of reverberation time, early decay time, clarity (C80), and center time (Ts) were taken according to the ISO 3382 standard. This data was compared to data from the great one-room concert halls. This comparison demonstrates that Bass Hall can function as an equally excellent concert hall, in spite of being a multi-purpose hall. [Research supported by the Bass Foundation.]

Contributed Papers

9:35

1aAA4. Consideration of the propagation loss of the direct sound with change in seat conditions. Yasushi Shimizu (Sonics in Architecture, Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180, shimiy@rpi.edu)

Reviewing excess attenuation of the direct sound by seats in concert hall acoustics gives the propagation loss by the performers absorption on the stage platform and diffraction of the front seats as well as seat dip effect in low frequency spectrum. Scale model experiments with 1/10 scale model in anechoic chamber were conducted at YAMAHA Acoustic Research Laboratories, Hamamatsu, focusing on propagation loss by (1) diffraction due to various geometrical conditions of the seating row in front of the listening position and (2) attenuation by the seat rows in front of the listening point. The results show the excess attenuation in middle and high frequency spectra. The optimum height between two adjacent seating rows is more than 1.3 feet and the number of seat rows are less than five seats. Normalized SPL referenced to SPL measured at 6.6 feet over the listening point is presented to be effective for evaluating excess attenuation. Concert Hall at Niigata Performing Arts Center was designed with consideration of these scale model experiments. Measurement data of normalized SPL and power spectra of orchestra sound in this hall are presented.

9:50

1aAA5. Measurement surveys of eight British concert halls. Russell Richardson and Bridget M. Shield (School of Eng. Systems and Design, South Bank Univ., London SE1 0AA, UK, shieldbm@sbu.ac.uk)

Objective measurement surveys and subjective surveys have been carried out in eight leading concert venues in the United Kingdom. In each hall a questionnaire survey of the whole audience at a concert of symphonic music was used to elicit the opinions of nonexpert listeners. Altogether over 3000 responses have been received. An objective survey of each (empty) hall has also been carried out, the parameters measured including reverberation time, early decay time, clarity index and definition. The number of measurement positions in each hall varied from 36 to 92. The questionnaire response data is to be correlated with the results of the objective measurements to investigate the relationship between objective data and the subjective impressions of nonexpert listeners. This paper will present summary data of the measurements in the eight halls and examine the variation in the values of the different parameters across each hall. The differences between the halls will also be considered.

10:05–10:15   Break

10:15

1aAA6. A five-channel microphone system for detecting 3-D acoustic image sources. Chulmin Choi, Lai-Hoon Kim (Seoul Natl. Univ., Seoul 151-742, South Korea), Se-Jin Doo (Dong-Ah Broadcasting College, Ans" Ansan, Kyonggi, South Korea), Yang-Ki Oh (Mokpo Natl. Univ., Mokpo, Chonnam 534-729, South Korea), Dae-Up Jeong (Chonbuk Natl. Univ., Jeonju, Chonbuk 561-756, South Korea), and Koeng-Mo Sung (Seoul Natl. Univ., Seoul 151-742, South Korea)

Measuring three-dimensional behavior of early reflections in a sound field has been an important issue in auditorium acoustics, since they are found to be strongly correlated with the subjective experience of spatial impression in rooms. A conventional way of detecting the direction and relative amplitude of reflections is to utilize a four-channel microphone system. This system, however, has the inherent possibility of missing some important reflections while exploring the early part of reflections. A new measurement system, which has five microphones on the apex of a tetrahedron and at the center of gravity, is proposed in the present work. The proposed system enables us to find more early reflections accurately, which previous four-channel microphone systems might miss. A peak detecting pair matching algorithm, which gives dominant peaks of impulse response automatically by typical iteration, was adopted, since we should find exact peak positions of measured impulse response in order to calculate source positions accurately. In the present paper, the theoretical backgrounds and features of the five-channel microphone system are presented. Also, some results from experiments using the system are discussed. Experimental results clearly show the advantage of the five-channel microphone system over the conventional four-channel microphone system. [Work supported by Korean Research Foundation Grant KRF-1999-1-310-004-3.]

10:30

1aAA7. Interaural cross-correlation: Sources of variability in concert halls. Gary S. Madaras (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, gary@talaske.com)

Impulse responses were measured inside multiple concert halls. Data analysis identified the level of the direct sound relative to the level of the architectural reflections as a significant source of IACC80 variability. Excluding the direct sound from the IACC80 integral durations (IACC80 × DIR) significantly decreased the variability. The remaining variability seemed to relate to the arrival direction of the architectural reflections. A testing model capable of various configurations and a scale measurement method were developed. Models that had reflections approaching the receiver from the sides produced much lower IACC80×DIR values than those without reflections approaching the receiver from the sides. The various configurations of architectural elements (i.e., balconies and stage canopy) and surface treatments (i.e., diffusive and specular) affected IACC80×DIR values to a lesser extent than altering the arrival direction of the architectural reflections. It was concluded that the general direction from which the architectural reflections arrive and the finish of the architectural surfaces are both sources of IACC80×DIR variability (the former having a much greater effect). It was also concluded that unless the direct sound is excluded from the integral duration, variability of interaural cross-correlation values does not relate as strongly to the arrival direction of the architectural reflections. [Work supported by the Concert Hall Research Group and the University of Florida.]

10:45

1aAA8. Investigation of early-to-late sound reflection effect in a room with a dome. Mokhtar Harun, Tharek Abdul Rahman, Md. Najib Ibrahim, and Ahmad Khan Said (Universiti Teknologi Malaysia, 81310 Skudai Johor, Malaysia, mokhtar@suria.fke.utm.my)

The material of a dome is normally very sound reflective. This is due to the fact that the material for inner surfaces of the dome is made of hard material for endurance and easy maintenance. Also, the dome is often situated high above the floor of the room with the intention of providing larger volume and thus creating comfortable envelopment of the dome for people inside it. Due to the height from the floor and the hard material of its inner surfaces, the dome naturally manifests strong sound reflection, sound focusing effects and uneven sound pressure levels. These phenomena also give rise to the reverberation time in a room with a dome. It is the intention of this paper to discuss the effects of early-to-late sound reflection with respect to reverberation time and sound reverberant level in a room with a dome.

11:00

1aAA9. Perceived spatial impression from various musical motifs.
Lily M. Wang (Arch. Eng. Prog., Univ. of Nebraska–Lincoln, 200B PKI, Omaha, NE 68182-0681, lwang@unl.edu)

Most of the objective measures which are currently used to quantify spatial impression in auditoria, such as lateral energy fraction (LF), interaural cross-correlation coefficient (IACC) and late lateral relative sound level (GLL), are calculated from an impulse response. For subjective testing, these impulse responses are often later convolved with music recorded in an anechoic environment to simulate music performance in the hall. One spatial impression measure, though, interaural level fluctuations (IALFs) which was first proposed by Griesinger [Aud. Eng. Soc., Preprint No. 3292 (1992)] and further developed by L. M. Wang and A. C. Gade [J. Acoust. Soc. Am. **107**, 2891(A) (2000)], is calculated from the musical source signal instead of from the impulse response. Preliminary investigations showed that the IALF, calculated from simulated sound fields generated by ODEON, changed relatively the same for different pieces of music across various halls. If IALF is a valid measure for spatial impression, then this indicates that the musical source is not critical in evaluating the difference in spatial impression between halls. In this paper, results from subjective listening tests are presented to determine if the same trend is perceived by listeners across different musical motifs, which would aid in validating IALF.

11:15

1aAA10. An experiment in the use of resilient materials in orchestra riser design. Ian B. Hoffman, Richard H. Talaske, and Byron W. Harrison (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, ian@talaske.com)

An experiment was assembled under extremely limited time and space constraints to investigate the acoustic effect of resilient materials in orchestra riser design. The riser was a portable, box-type riser for assembly on a flat concert platform, not a lift-type riser. The primary goal of the inquiry was to maximize lateral vibration transfer through the riser (from celli, basses, etc.) without detrimentally affecting airborne sound radiation from the riser. A portion of a full-scale riser was mocked up, including sections of differing elevation. Various design configurations of the riser allowed experimentation with and without resilient materials, and with resilient materials placed at different positions within the riser structure. The experiment was conducted within a reverberant volume. The riser was excited by mechanical shakers and by a cello, independently. Vibration measurements were captured using accelerometers on the riser, and airborne sound pressure levels were captured by a distant microphone. Initial results suggest that the use of resilient materials in certain configurations can considerably increase lateral vibration transfer through the riser (as compared to rigid construction), without significant compromise to airborne radiation from the riser.

MONDAY MORNING, 4 JUNE 2001   ADAMS ROOM, 8:30 TO 11:55 A.M.

Session 1aAO

Acoustical Oceanography: Acoustical Instrumentation for Water Column Measurements I

Kenneth G. Foote, Chair

Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

Chair's Introduction—8:30

Invited Papers

8:35

1aAO1. Multibeam sonars: Applications for fisheries research. Larry A. Mayer, Yanchao Li (Ctr. for Coastal and Ocean Mapping, Univ. of New Hampshire, Durham, NH 03824, lmayer@unh.edu), and Gary Melvin (Dept. of Fisheries and Oceans, St. Andrews, NB, Canada)

Multibeam sonars are rapidly becoming a standard tool for seafloor mapping in support of geological, geophysical, and engineering applications. More recently, the ability of multibeam sonars to provide high-resolution, large areal coverage, and potentially quantitative, coregistered, backscatter has been applied very successfully to problems of defining fisheries habitat. While most multibeam sonars are designed to gate out all midwater returns, newly developed systems now allow access to the full data stream and thus offer the possibility of application to studies of pelagic and demersal fisheries. Traditional acoustic approaches to fisheries issues have used single beam echo sounders that sample a relatively small volume of the water column within a survey area. Multibeam sonars provide a mechanism to greatly enhance both the resolution and the area of coverage. When combined with powerful new 3-D visualization techniques, they can offer immediate feedback on fish behavior as well as the critical question of vessel avoidance. If properly calibrated, multibeam sonars can provide the means for much more robust assessment of stock levels and perhaps even species identification.

9:35

1aAO2. Physical constraints in sonar design. Philip R. Atkins and Tim Collins (School of Electron., Univ. of Birmingham, Edgbaston, Birmingham B15 2TT, UK, p.r.atkins@bham.ac.uk)

The dynamic range and processing gains achievable within a simple passive sonar will be reviewed. It will be demonstrated that the constraints of modern electronic hardware components do not necessarily limit the performance of a passive sonar system when compared with the physical constraints of the acoustic channel. By contrast it will be demonstrated that the pulse design of an active sonar places significant limitations on the capability of the system. For example, range resolution, Doppler resolution, range sidelobes, ambiguities and reverberation performance all depend on the design of the sonar transmission. Techniques available to the designer

---

for suppressing range sidelobes and pulse designs with both range and Doppler sensitivity will be reviewed. The constraints imposed by the transducer will be considered, particularly in relationship to ultrahigh range resolution capabilities. Finally, a comparison between an early 1960s sonar and a modern equivalent will be made in terms of performance, power consumption and size.

10:05–10:25   Break

Contributed Papers

10:25

1aAO3. Acoustic classification of individual zooplankton using artificial neural network. Andone C. Lavery and Timothy K. Stanton (Woods Hole Oceanogr. Inst., Dept. of Appl. Ocean Phys. and Eng., Woods Hole, MA 02543)

The acoustic scattering characteristics of zooplankton generally fall into one of three categories: gas-bearing, fluidlike, or elastic shell. Acoustic scattering models are available for zooplankton that fall into each of these categories. A multilayered feedforward artificial neural network, based on a backpropagation algorithm, has been used to discriminate between the model-based acoustic scattering responses of individual zooplankton from these different zooplankton classes. Realistic animal size distributions, taken from net tows, have been used as input to the models to create (simulated) acoustic data for use as input into the classification network. Two different approaches have been taken to selecting input features from the model-based zooplankton scattering responses, and the results of the neural network classification for these two features sets are compared. The feasibility of applying this neural-network-based classification technique to both multifrequency and broadband target strength data from field surveys is discussed.

10:40

1aAO4. Simple, nonoptimal, real-time classifiers for euphausiid and copepod numerical density estimation. Philip R. Atkins, Claire Bongiovanni, David T. I. Francis (School of Electron., Univ. of Birmingham, Edgbaston, Birmingham B15 2TT, UK), Kenneth G. Foote (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Tor Knutsen (Inst. of Marine Res., N-5817 Bergen, Norway), Peter K. Eriksen, Mette Torp Larsen, and Tom Mortensen (RESON A/S, DK-3550 Slangerup, Denmark)

An active sonar covering the frequency range 1.6–3.2 MHz was used to observe zooplankton at two sites along the Norwegian coast during the period 28 April–9 May 1999. The sonar transmitted LFM signals with an approximate bandwidth-time product of 100. The matched-filter envelope was computed within the receiver at a sampling rate of 10 M samples/s. Freshly caught, swimming specimens were insonified *ex situ* in a tank mounted on the stern of R/V JOHAN HJORT, and the envelope signal was classified by eye as belonging to the euphausiid *Meganyctiphanes norvegica*, or copepod *Calanus finmarchicus*. The separated signals were then averaged to produce templates corresponding to the two classes. All orientation information associated with the specimens was thus removed. A sliding-window-normalized covariance was calculated between the templates and the receiver output for a large number of *ex situ* and *in situ* pings, a technique capable of being implemented at the sampling rate of the receiver. Although simple, this multiple-hypothesis-testing procedure outperformed a conventional split-window (CFAR) normalizer and produced a rapid and apparently reliable method of distinguishing the two classes and estimating their numerical density. [Work supported by EU RTD Contract No. MAS3-CT95-0031.]

10:55

1aAO5. Observation of fish length and behavior of an individual fish using echo trace analysis. Kouichi Sawada, Yoichi Miyanohana, Hideyuki Takahashi, and Yoshimi Takao (Natl. Res. Inst. Fisheries Eng., Fisheries Res. Agency, Hasaki, Kashima, Ibaraki 314-0421, Japan, ksawada@nrife.affrc.go.jp)

A compact sized split-beam echo sounding system was developed and used in Sagami Bay, Japan, to validate echo trace analysis (ETA) that can observe the length and the behavior of an individual fish. Fishing by hook and line was conducted to confirm the fish species and its length distribution. The echo sounding system was tethered from an anchored ship. Electric power was provided through a cable. This system consists of a split-beam transducer (70 kHz, 118 mm), a transmitting and receiving unit, a data-recording unit, and a control unit and the total weight of the echo sounding system is about 30 kg. Recorded data are up linked to the computer on the ship through a LAN cable. The fish length and fish swimming speed were estimated using ETA. From the fishing results, 79l (*Scomber japonicus*) or horse mackerel (*Trachurus japonicus*) had an average fork length (FL) of 27.9 cm. The estimated fish length distribution coincides with that obtained from a fishing well. The estimated swimming speed was 0–1.2 m/s and corresponds to 0–4.3 (FL). These results show the validity of ETA.

11:10

1aAO6. Use of multiple-beam echo-sounders for estimation of target strength *in situ*. Ian Hampton (BENEFIT, P.O. Box 912, Swakopmund, Namibia) and Michael A. Soule (Marine and Coastal Management, Dept. of Environ. Affairs and Tourism, Roggebaai, 8012, Cape Town, South Africa)

In principle, the average target strength of fish in a defined volume can be estimated *in situ* from the ratio between the volume backscattering strength and a count of the number of fish in the volume. The chief problems, which have precluded use of the method in practice, are the difficulty in obtaining reliable counts at common fish densities and in defining beam volume. It is suggested that both problems could be solved to a large extent by using a relatively modest high-frequency, multi-beam echo-sounder as a counter. This presentation outlines the underlying principle of the method, considers the design criteria, and describes a 48-element 420 kHz system, (code named ABACUS) which is being built to test the method. The system, which has been designed to operate to approximately 70 m, for target strengths in the range 40 to 55 dB, generates a single transmit beam and 32 receiving beams orthogonal to it, which intersect to give up to ten 1 deg × 1 deg resolution cells. Preliminary test-tank results are presented, and plans for further experimental work discussed. It is believed that the method should be capable of delivering *in situ* target strength estimates at densities well beyond the limit of current methods.

11:25–11:55

Panel Discussion

Session 1aPA

Physical Acoustics: Acoustic Detection of Land Mines and Other Buried Objects

James M. Sabatier, Chair

National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Contributed Papers

10:00

1aPA1. Acoustic landmine detection at 40° below zero. James M. Sabatier, Ning Xiang, and George Nock (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, MS 38677)

To study the effects of snow cover and frozen ground, mine detection experiments based on acoustic to seismic (A/S) coupling have been conducted on a roadway at the Defense Research Establishment Suffield (DRES), Alberta, Canada, 5–16 February 2001. The site is a gravel road with grain or gravel sizes as large as several centimeters. The ground was frozen. Only a few measurements of acoustic-to-seismic coupling in frozen grounds have been reported. The coupling ratio is within a factor of 10 of that for unfrozen soils. The fast compressional wave speeds were expected to be significantly larger than those in unfrozen soils, while slow compressional wave speeds were expected to be possibly unchanged, except for temperature effects. Therefore, one might expect significantly larger wavelengths for the fast compressional wave and nonlocally reacting coupling. Compensation for these changes included using higher-frequency sound and normal incidence angles for the sound source. This presentation outlines the results of this testing. [Work supported by the U.S. Army Communcations–Electronics Command Night Vision and Electronic Sensors Directorate.]

10:15

1aPA2. Acoustic scattering by a buried object with a compliant top in a rigid porous material. Doru Velea, Roger Waxler, James M. Sabatier, and Craig J. Hickey (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677)

Recently, a method for the detection of landmines buried a few inches below the surface of the ground has been developed at the University of Mississippi. This method utilizes acoustic-to-seismic coupling and measures the particle velocity of the insonified ground surface on- and off-target with a Laser Doppler Vibrometer. As part of the on-going theoretical effort to explain the measurements, this paper is concerned with acoustical scattering from an object buried in a rigid porous half-space. The object is assumed to be a rigid right circular cylinder with the exception that its top can move in the vertical direction in a pistonlike fashion with a mechanical resonance at around 100 Hz. A boundary element method has been employed to predict the on-target/off-target particle velocity ratio. Comparisons with observations are made. [Work supported by the U.S. Army Communications-Electronics Command Night Vision and Electronic Sensors Directorate.]

10:30

1aPA3. Measurement of attenuation and speed of sound in soils for the purposes of imaging buried objects. Michael Oelze, Robert Darmody, and William O'Brien (Univ. of Illinois at Urbana–Champaign, 405 N. Mathews, Urbana, IL 61801, oelze@brl.uiuc.edu)

The speed and attenuation of sound vary in different soil types and with different levels of water saturation. Imaging techniques make use of speed and attenuation to range targets and compensate for loss in echo magnitude with depth. Signals from an acoustic source over the range of 2–6 kHz were passed through soil and detected by a coupled hydrophone. Six soil samples were used to represent a range of properties expected to influence acoustic response. Clay ranged from 2%–38%, silt from 1%–82%, sand from 2%–97%, and organic matter from 0.1%–11.7%. The attenuation and propagation speed were determined in the soil as a function of soil moisture and compaction. Attenuation coefficients ranged from 0.12–0.96 dB cm^{-1} kHz^{-1}. Correlation coefficients were 0.35 ($P=0.01$) and 0.31 ($P=0.03$) between attenuation and soil water content and soil bulk density. Propagation speeds ranged from 86 to 260 m s^{-1}. The correlation coefficient with speed was -0.28 ($P=0.05$) for soil water content and 0.42 ($P=0.002$) for porosity. Given the attenuation, it is theoretically possible to image objects in soils to 40 cm in depth. [Work supported by U.S. Army Construction Engineering Research Laboratory, Champaign, IL.]

10:45

1aPA4. Acoustic landmine detection: Poroelastic model and elastic wave models. Yanqing Zeng and Qinghuo Liu (Dept. of Elec. and Computer Eng., Duke Univ., Box 90291, Durham, NC 27708-0291, qhliu@ee.duke.edu)

Acoustic waves can be a viable tool for the detection and identification of landmines and unexploded ordnance (UXO). Design of acoustic instruments and interpretation and processing of acoustic measurements require accurate numerical models for acoustic wave propagation in heterogeneous soil with buried objects. Two numerical models will be compared in this work: the soil is treated as an elastic material in the first model, and as a poroelastic material in the second model. A multidomain pseudospectral time-domain method is used for the elastic model, while a finite-difference method is used to solve Biot equations in the poroelastic model. Acoustic wave response of a buried landmine is investigated with both methods. When the soil is treated as a fluid-saturated poroelastic medium, additional attenuation is automatically introduced via the relative motion between solid frame and pore fluid. Thus the poroelastic model is presumably a more accurate model; however, the elastic model requires less computation time. This work will study several important scenarios such as landmines buried in dry sand, in wet sand, and in mud. A comparison between results of the PSTD model and results of the poroelastic model is performed to characterize the regime of validity for the elastic model.

11:00

1aPA5. A hybrid numerical model to predict the effect of buried objects on the acoustic-to-seismic coupling on ground surfaces. Shahram Taherzadeh[a] (Natl. Ctr. for Phys. Acoust., Coliseum Dr., Univ. of Mississippi, Oxford, MS 38677, s.taherzadeh@open.ac.uk)

A numerical model based on a hybrid boundary integral equation/wave number integration is presented and is used to predict the influence of rigid objects buried within a layered, porous elastic soil upon excitation of the ground surface induced by an airborne sound source. Recently, acoustic-to-seismic coupling phenomena have been used successfully to detect buried landmines [J. Sabatier and N. Xiang, J. Acoust. Soc. Am. **106**, 2143 (1999), **107**, 2896 (2000)]. If a ground surface is subjected to insonification, the sound wave penetrates the ground and is reflected and refracted from the topsoil layers, resulting in complex acoustic-to-seismic spectra. The presence of an object alters these spectra by scattering and/or enhanc-

ing certain vibrational modes. The model will be applied to predict these changes in the presence of buried landmines. [a] Permanent address: Faculty of Technology, The Open Univ., Milton Keynes, UK.

11:15

1aPA6. 3-D multidomain PSTD for elasticity. Yanqing Zeng and Qinghuo Liu (Dept. of Elec. and Computer Eng., Duke Univ., Box 90291, Durham, NC 27708-0291, qhliu@ee.duke.edu)

A 3-D numerical modeling scheme is developed in general curvilinear coordinates by using a multidomain pseudospectral time-domain (PSTD) method. The computational domain is naturally divided into subdomains of piecewise homogeneous, curvilinear quadrilaterals conforming to the problem geometry. These curvilinear quadrilaterals are then mapped onto a unit square, where the spatial derivatives are approximated by the Chebyshev PSTD method. Characteristic conditions and physical boundary conditions are used to patch the solution between adjacent subdomains. A newly developed unsplit perfectly matched layer (PML) is used to absorb outgoing waves at the computational edge. This multidomain PSTD algorithm allows natural incorporation of curved boundaries into the numerical scheme, thus providing more increased flexibility and accuracy than the conventional finite-difference method. The numerical scheme is validated by analytical solutions and by other numerical methods. Applications of this multidomain PSTD method will be shown for acoustic landmine detection.

MONDAY MORNING, 4 JUNE 2001

STATE BALLROOM, 9:00 A.M. TO 12:00 NOON

Session 1aPP

Psychological and Physiological Acoustics: Perception, Learning, and Cognition

Beverly A. Wright, Chair

Audiology and Hearing Science Program, Northwestern University, 2229 North Campus Drive, Evanston, Illinois 60208

Contributed Papers

9:00

1aPP1. The perceptual interaction between the pitch and timbre of musical sound. Jeremy P. Marozeau, Alain de Cheveigné, Stephen McAdams, and Suzanne Winsberg (IRCAM-CNRS, 1 place Igor-Stravinsky, F-75004 Paris, France)

The perceptual dependency of the timbre of musical sounds on their fundamental frequency ($F0$) was examined through two timbral dissimilarity experiments. In Experiment 1, subjects judged timbral dissimilarity between pairs of stimuli produced by 12 musical instruments equal in $F0$, duration, and loudness. There were three sessions, each at a different $F0$ (B2, Db3, Bb3). In Experiment 2, the same stimuli were presented in pairs with a constant difference in $F0$. Two sessions corresponded to two different intervals (B2-Db3 and B2-Bb3). Subjects were instructed to ignore pitch difference. The results are interpreted in terms of an evolution of timbre space as a function of $F0$. Experiment 1 produced dissimilarity matrices that were similar at different $F0$s, suggesting that the relative positions of instruments within the timbre space varied little with $F0$. Experiment 2 produced dissimilarity matrices that were symmetrical, suggesting further that $F0$-dependent displacements within the timbre space are small. Patterns of dissimilarity were similar with or without an $F0$ difference, showing that subjects could ignore a constant difference in pitch when making timbral dissimilarity judgments. Results suggest that the perception of timbre is relatively independent of pitch and can be abstracted across differences in $F0$ under these conditions.

9:15

1aPP2. Infants' perception of transposed melodies. Daniella A. Kim and Lynne A. Werner (Dept. of Speech and Hearing Sci., Univ. of Washington, 1417 NE 42nd St., Seattle, WA 98105)

It has been suggested that infants are similar to adults in their perception of tone chroma, in the perceptual similarity of notes in octave relationships [Demany and Armand (1984)]. In music, note pitches are defined by height (frequency distance) and chroma (key). The current study further examines this question. Seven- to nine-month-old infants' perceptions of transpositions of six-note tonal melodies were measured using an observer-based procedure. Once given a background melody subjects were trained to respond to an entirely different training melody and not to a simple transposition. Once that discrimination had been learned, transposed versions of the background melody were introduced. If infants perceive tone chroma as adults do, they should learn this discrimination more quickly for an octave transposition. In fact, infants learned to ignore a transposition by a fifth more quickly than transpositions by an octave. The implication is that infants found melodies that were transposed by a smaller distance in frequency as more perceptually similar to the background melody than they did for the octave transposition. A possible conclusion is that octave transpositions of whole melodies have special perceptual status for infants of this age range. [Work supported by NIDCD Grant No. DC00396.]

9:30

1aPP3. Auditory streaming without spectral cues. Nicolas Grimault, Sid P. Bacon (Psychoacoustic Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., Tempe, AZ 85287-1908), and Christophe Micheyl (UPRESA CNRS 5020, Pav. U, Hp. E. Herriot, Lyon 69437, France)

Spectral features are powerful cues in the perceptual organization of sound sequences. The ''peripheral channeling'' explanation of stream segregation predicts that the perception of fusion or fission in sequences, in a repeating ABA format, depends on the degree of spectral overlap of the A and B sounds. Nevertheless, previous studies have shown that other factors can also affect auditory stream segregation. Sequences of fully unresolved complex tones, bandpass filtered in the same frequency region, have been shown to perceptually split apart into two streams when the fundamental frequencies of the complexes are widely separated [Grimault *et al.*, J. Acoust. Soc. Am. **108**, 263–271 (2000)]. In the present experiment, normally hearing adults heard rapid alternations of a pair of flat-spectrum, sinusoidally amplitude-modulated broadband noise bursts. These noise bursts were segregated based on differences in modulation rate, as determined with a constant stimulus procedure as well as with a procedure in which segregation responses were collected continuously while the modulation rate of B was swept up or down during a block of trials. [Work supported by Fondation pour la Recherche medicale and NIDCD.]

2288   J. Acoust. Soc. Am., Vol. 109, No. 5, Pt. 2, May 2001

141st Meeting: Acoustical Society of America   **2288**

9:45

1aPP4. Loudness-independent cues to object striking force. Stephen Lakatos (Dept. of Psych., Washington State Univ., 14204 NE Salmon Creek Ave., Vancouver, WA 98686) and James W. Beauchamp (School of Music, Univ. of Illinois–Urbana-Champaign, Urbana, IL 61801)

To what degree can listeners rely on auditory cues besides loudness to correctly judge the striking forces used to excite an object? Listening experiments used a discrimination task to assess whether individuals could correctly map loudness relations onto striking forces for stimuli derived from tam–tam, tympani, and xylophone-bar recordings. On each trial, listeners heard two sound pairs representing an instrument struck with differing strike forces using a mallet; one pair had its loudness relations intact while the other had its loudness relations inverted. Listeners determined which pair preserved its correct loudness relations. Both striking force and mallet properties served as independent variables. Discrimination varied in general with the extent to which the mallet and striking force excited the characteristic vibrational modes of the instruments: the more acoustic modes available, the higher the discrimination performance. Discrimination was significantly poorer for instruments with fewer acoustic cues generated by such modes, like the xylophone, than it was for instruments with a rich set of cues, like the tam–tam. Mallet density and weight were also important factors in modulating discrimination. The results indicate source-based cues besides loudness can convey information about striking force, albeit with less effectiveness. [Research supported by Air Force Grant F49620-99-1-0293.]

10:00

1aPP5. Perceptual learning and attentional cues in nonsimultaneous masking. Marina M. Rose, Douglas E. H. Hartley, and David R. Moore (Univ. Lab. of Physiol., Oxford Univ., Parks Rd., Oxford OX1 3PT, UK, marina.rose@physiol.ox.ac.uk)

This study investigated training effects in three nonsimultaneous masking conditions, and the possible role of an attentional cue. Three groups of six participants trained on seven consecutive days on either (i) backward masking (BM, 0-ms delay), (ii) combined forward and backward masking with a short inter-masker delay (10-ms, FBS) or (iii) combined forward and backward masking with a long inter-masker delay (110-ms, FBL). The signal was a 10-ms, 1-kHz tone pip, and the masker was a 300-ms, 1.6-kHz wide noise (40 dBSPL/Hz), centered at 1 kHz. Training sessions had 15 runs (30 presentations/run), using a 2AFC, maximum-likelihood procedure. The FBS condition showed the highest amount of masking, followed by the BM and FBL conditions, respectively (p <0.05). A small but significant improvement was found between training days when all three groups were combined (p<0.001). Pre-training data showed, for most subjects, lower FBL than BM thresholds. This could indicate that the first noise burst in the FBL condition did not mask the signal, but rather acted as a temporal and attentional cue, helping to ''listen at the right time.'' [Work supported by the National Health Service, the Wellcome Trust, and the Medical Research Council.]

10:15

1aPP6. Rapid improvements on interaural-time-difference discrimination: Evidence for three types of learning. Jeanette A. Ortiz, Beverly A. Wright, Matthew B. Fitzgerald, and Jagan Pillai (Audiol. and Hearing Sci. Prog., 2299 N. Campus Dr., Northwestern Univ., Evanston, IL 60208-3550, j-ortiz@northwestern.edu)

Listeners often show rapid improvements when first performing auditory discrimination tasks. This learning was explored by investigating how training 300–1500 trials on one of three conditions (all 2IFC) influenced interaural-time-difference (ITD) discrimination on a target ITD condition (lateralization, 0-μs ITD, 500 Hz). Listeners trained on ITD or interaural-level-difference (ILD) discrimination had equivalent ITD-discrimination thresholds that were significantly lower than those of naive listeners. However, listeners trained on duration discrimination had ITD-discrimination thresholds between those of naive and ITD-trained listeners. These data suggest that rapid decreases in ITD-discrimination threshold result partly from learning the general procedure (induced by the duration training) and

partly from learning the lateralization task (induced by the ILD training). In contrast, compared to naive listeners, final within-subject standard deviations of the ITD-discrimination thresholds were significantly smaller for ITD-trained, somewhat smaller for ILD-trained, and no smaller for duration-trained listeners. These data suggest that rapid reductions in threshold variabiilty result partly from learning the lateralization task (induced by the ILD training) and partly from learning something provided, here, only by ITD training. Thus, at least three different types of learning can contribute to initial improvements in performance on auditory discrimination tasks. [Work supported by NIDCD.]

10:30–10:45   Break

10:45

1aPP7. The impact of labeling versus detection training on performance in an asynchrony-detection task. Julia A. Mossbridge and Beverly A. Wright (Audiol. and Hearing Sci. Prog., 2299 N. Campus Dr., Northwestern Univ., Evanston, IL 60208-3550, j-mossbridge@northwestern.edu)

It is not clear which training methods yield the fastest improvement on perceptual skills. To address this issue, performance on a detection task was compared following training on a labeling or detection task. The labeling task was to determine, for each presentation, whether the onsets of two tones were synchronous or asynchronous. Tone onsets were separated by one of six durations. The detection task was to determine which of two presentations contained tones with asynchronous onsets. The duration between tone onsets was varied adaptively to establish the asynchrony-detection threshold (79% correct detections). In both tasks, the tonal frequencies were 0.25 and 4 kHz. Labeling-trained listeners (n =11) had detection thresholds lower than those of untrained controls (n =6) following just 600 labeling trials (p =0.008). In contrast, detection-trained listeners (n=6) had detection thresholds lower than controls only following 3660 detection trials (p=0.027), 2160 trials on the asynchrony-detection task plus 1500 trials on related temporal-acuity tasks. Thus, labeling training appears to improve performance on an asynchrony-detection task more quickly than detection training itself. By increasing our understanding of perceptual learning processes, these data may aid the treatment of individuals with disorders in auditory perception. [Work supported by NIDCD.]

11:00

1aPP8. Specificity of learning in an auditory asynchrony-detection task. Beverly A. Wright, Matthew B. Fitzgerald, and Julia A. Mossbridge (Audiol. and Hearing Sci. Prog., 2299 N. Campus Dr., Northwestern Univ., Evanston, IL 60208-3550)

The ability to determine whether the onsets of two tones at different frequencies are asynchronous is a measure of temporal acuity. Learning in this asynchrony-detection task and its generalization to untrained frequencies and temporal-acuity tasks was explored. For six listeners, asynchrony-detection thresholds for tones at 0.25 and 4 kHz improved from 79 to 20 ms with 4320 training trials. A z-score analysis revealed that trained listeners learned significantly more than untrained controls on the trained condition, but only slightly or no more than controls on: (1) the trained task at two sets of untrained frequencies (0.5 and 1.5 kHz, 0.75 and 1.25 kHz), and (2) three untrained temporal-acuity tasks at the trained frequencies (offset asynchrony detection, onset temporal order, and offset temporal order). This generalization pattern parallels that seen when the trained task was to determine the onset order of two tones [B. A. Wright and J. A. Mossbridge, J. Acoust. Soc. Am. **107**, 2882 (2000)]. Thus, learning in both asynchrony-detection and temporal-order tasks appears to be mediated by task-specific, frequency-dependent mechanisms. By increasing our understanding of temporal-acuity mechanisms and their plasticity, these data may guide the treatment of related hearing disorders. [Work supported by NIDCD.]

11:15

1aPP9. Psychophysical studies of auditory sensory traces. Denis McKeown (School of Psych., Univ. of Leeds, Leeds, UK), David Wellsted (Univ. of York, York, UK), and Tim Green (Univ. College London, London, UK)

The evidence is impressive, using an evoked potentials paradigm, that detailed internal models of recent sound stimulation are formed and compared with incoming sounds: Changes are incorporated, it is supposed, in order to update those models or sensory traces. Here psychophysical studies are reported. The first experiment poses the question of how recent tone stimulation affects ongoing performance in a discrimination task. Briefly, it is assumed that such a tone, or cue, may provide a highly frequency-specific sensory trace equivalent to sensory enhancement at that frequency. In a discrimination task two complex sounds occur in successive intervals. A single component of one complex is decreased in intensity: The task is to report whether this is in the first or second interval. Now, prior to each trial a tone cue is introduced: This is at the same frequency as the decreased component, or at a different frequency. The reasoning is that if the cue tone produces sensory enhancement, then it should effectively remove the decrement when cue and decrement are at the same frequency, impairing performance on these trials. This was the result obtained. Other experiments ued a detection task in examining the time source of the presumed sensory trace.

11:30

1aPP10. Unsigned coding of stimulus change. Ervin R. Hafter, Erick Gallun (Dept. of Psych., Univ. of California, Berkeley, CA 94720), Anne-Marie Bonnel (Ctr. Natl. de la Recherche Scientifique, Marseille, France), Brent Edwards, and Neal Viemeister (Univ. of Minnesota, Minneapolis, MI 55455)

Various studies have pointed to a special importance for stimulus transients in such diverse areas as the overshoot in masking and onset dominance in sound localization. Evidence that the internal representation of transience may be unsigned is found in cases where listeners readily detect

a change in the ongoing stimulus without successfully identifying the direction of that change. We will review these results and present new data on sensitivity to change without sign in such dimensions as frequency and lateral location. Unusual sensitivity to transience has implications for the perception of natural scenes, where change per se may represent the code for simple alerting signal. [Work supported by grants from NIH, DC00087 and DC00683.]

11:45

1aPP11. Are cognitive processes involved in the perceived degree of urgency of auditory warning signals? Anne Guillaume, Lionel Pellieux, Veronique Chastres (Imassa BP73, 91223 Bretigny sur Orge, France), and Carolyn Drake (Universite de Paris V, Boulogne-Billancourt Cedex, France)

Studies on the perceived degree of urgency of auditory warning signals have demonstrated that the greatest sense of urgency is created by sequences with a high fundamental frequency, many irregular harmonics, and fast sequence rate (Edworthy *et al.*, 1991). However these effects disappeared when subjects were under high workload. In order to further investigate the cognitive processes involved, listeners' perception of two sets of warning sounds were tested: one constructed following Edworthy's indications, and a second composed of real alarms currently used by the French airforce. A first group of listeners evaluated the perceived degree of urgency between pairs of alarms. A second group quantified the perceived difference between the two alarms. A multidimensional analysis demonstrated that the spatial distribution of the *constructed alarms* was similar for both tasks. However, for the *real alarms*, a different pattern was observed: Some alarms were classified as nonurgent although their acoustical characteristics would have indicated a high-urgency classification. The conclusion is that the perceived degree of urgency is determined by both low-level perceptual factors and more complex cognitive processes involving long-term memory and personal experience.

MONDAY MORNING, 4 JUNE 2001   RED LACQUER ROOM, 8:30 A.M. TO 12:00 NOON

Session 1aSC

Speech Communication: Vowels, Sonorants, Consonants: Acoustics and Perception (Poster Session)

Amy T. Neel, Chair

Department of Audiology and Speech Sciences, Purdue University, 1353 Heavilon Hall, West Lafayette, Indiana 47906

Contributed Papers

All posters will be on display from 8:30 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 8:30 a.m. to 10:15 a.m. and contributors of even-numbered papers will be at their posters from 10:15 a.m. to 12:00 noon.

1aSC1. Effects of instructions to produce hyperarticulated vowels on formant values. Marianne Pouplier (Haskins Labs, 270 Crown St., New Haven, CT 06511 and Yale Univ., marianne.pouplier@yale.edu) and D. H. Whalen (Haskins Labs, New Haven, CT 06511)

Clear speech collected in the laboratory is typically well articulated, but speakers can produce even clearer speech (''hyperarticulated'') when so instructed. Johnson *et al.* [Language **69**, 505–528 (1993)] found that hyperarticulations, averaged across speakers, are more extreme in both the

$F1$ and $F2$ dimension. Here, 10 speakers of a Rhode Island dialect of English recorded words with 11 different vowels in (h)Vd environments (e.g., ''heed,'' ''aid''). Individuals show the effect for between 2 and 9 vowels; there was no vowel that had more extreme formant values (for $F1$ or $F2$) for all 10 speakers. (Averaged across talkers, only the low front vowel had more extreme values in both $F1$ and $F2$.) Typically, hyperarticulations that were not more extreme in formant values were equivalent to the normal production, while a few had less extreme values. The lack of consistency in the individual results calls into question the use averaged values, since speakers may be changing articulatory parameters in a more

complex way. The dynamic nature of some of the vowels may have an influence, since the initial measurements were taken from a single point in the syllable. [Work supported by NIH.]

1aSC2. On trading relations in vowel production: An x-ray microbeam study. Angela Slama and Gary Weismer (Univ. of Wisconsin–Madison, 1975 Willow Dr., Madison, WI 53706, slama@waisman.wisc.edu)

The purpose of this study was to investigate the articulatory and acoustic variability in vowel production, across 52 normal speakers of English. Specific to that variability, the notion of trading relations was addressed. In a study by Perkell *et al.* (1993), trading relations describes the complementary nature by which two different articulatory mechanisms are employed by a single speaker, when producing a given speech sound. Perkell and colleagues found weak, negative correlations within three out of four speakers, for the tongue-raising and lip-rounding mechanisms used to produce the English vowel /u/. These findings suggest that within a given speaker, the two primary articulatory mechanisms used to produce /u/ may be traded, or used in a complementary fashion, to constrain the acoustic variability of the vowel. In the current experiment, kinematic and acoustic measurements were used to explore the possibility of the same kinds of trading relations across speakers. The specific question asked was, do speakers who demonstrate large amounts of lip rounding for the vowel /u/, also use small amounts of tongue backing and/or rounding, and vice versa? Interspeaker variability for the acoustic and kinematic measures will be evaluated and discussed relative to this question. [Work supported by NIH No. DC03274.]

1aSC3. Male and female vowels identified by visual inspection of raw complex waveforms. Michael A. Stokes (MAS Enterprises, 3926 Graceland Ave., Indianapolis, IN 46208, masmodel@indy.net)

A perceptual experiment involving visual identification of the raw complex waveforms of American English vowels from two male speakers was presented at the 131st meeting of Acoustical Society of America (Stokes, 1996). In that study, a subject correctly identified five out of nine vowels for both male speakers using only a visual presentation of the raw complex waveforms. The present study replicates those earlier results with the same subject correctly identifying four out of nine vowels and six out of nine vowels for two new male speakers. Furthermore, four out of nine vowels produced by a female speaker were correctly identified. These results demonstrate that the visual cues being used for vowel identification can be applied across genders, in addition to extending the work from 1996 to now show successful visual identification of vowels across four male talkers using only raw complex waveform displays. These results will be discussed, including an analysis of the errors made by the subject. A description of the visual cues and the new model of vowel perception and production that resulted from these cues can be found at http://www.indy.net/~masmodel.

1aSC4. Exploring the "hyperspace" effect in vowel perception. D. H. Whalen (Haskins Labs, 270 Crown St., New Haven, CT 06511) and Harriet Magen (Haskins Labs, New Haven, CT 06511)

When using the method of adjustment to choose the best representative of an English vowel in an $F1$ by $F2$ space, listeners have been found to select formant values more extreme than those they actually produce [Johnson *et al.*, Language **69**, 505–528 (1993)]. This preference for "hyperspace" targets (those beyond actual productions) is argued to indicate a phonetic target more extreme than the subjects' actual hyperarticulations. The same methodology was used with 10 Rhode Island subjects. They produced vowels under two conditions, citation and then (after the perceptual task) hyperarticulated speech. Perception was assessed with a grid of 298 $F1/F2$ combinations from which listeners chose the best match for a keyword's vowel. Females and males were nearly equivalent in percep-

tion, as assumed previously but not shown. However, only one vowel was clearly extreme in both formants, with seven other vowels being more extreme on one formant, two vowels showing hypoarticulation, and the final one showing no effect. This lack of a robust effect may be due to difficulty of using $F2$ in perceiving the synthetic vowels. It may also be that the listeners were selecting perceptual targets from a more neutral dialetic than their own. [Work supported by NIH.]

1aSC5. Variations in coda sonorants at different speaking rates. Jodi P. Bray (Univ. of Florida Prog. in Linguist., P.O. Box 115454, Gainesville, FL 32611-5454)

Despite sonority's role in phonological accounts of syllable structures in the world's languages, phonetic correlates of sonority remain controversial. This study is part of a larger project designed to examine possible phonetic cues to sonority. Specifically, this paper reports on acoustic characteristics (i.e., duration, amplitude, and $F0$) of sonorants [r, l, m, n] occurring word finally or before an obstruent. The goal is to examine (a) whether these acoustic characteristics vary across speaking rate: slow, casual, and rapid; (b) whether these characteristics vary across speakers. Preliminary results based on data from one male native English speaker suggested that sonorant duration did not seem to vary at different speaking rates. Moreover, the ratio of sonorant to rhyme duration is greatest in rapid speech suggesting that sonorants require duration for their perceptibility. Additional speakers will be analyzed to confirm these findings. A perception experiment designed to examine degree of perceptibility of sonorants at different speaking rates and in different degrees of noise will also be conducted. Perceptual data will be examined and discussed in relation to production data.

1aSC6. A documentation of phonetic detail in onset-embedded words. Katherine Crosswhite (Dept. of Linguist., Univ. of Rochester, Rochester, NY 14627, crosswhi@ling.rochester.edu), Mikhail Masharov (Dept. of Brain and Cognit. Sci., Univ. of Rochester, Rochester, NY 14627, mym@bcs.rochester.edu), Joyce M. McDonough (Univ. of Rochester, Rochester, NY 14627, jmmcd@ling.rochester.edu), and Michael Tanenhaus (Univ. of Rochester, Rochester, NY 14627, mtan@bcs.rochester.edu)

As part of a study of online processing of onset-embedded words, we documented phonetic-level properties of speech production, such as duration differences, that may cue for word identification, segmentation and/or lexical access. Forty nouns (20 pairs like cap/captain, doll/dolphin) were placed in sentences in three frames (nuclear, prenuclear, and utterance-final position). Each sentence was read, then pronounced from memory by eight speakers randomly ordered. In keeping with previous research, instrumental analyses revealed strong effects of syllable count on vowel duration [Klatt, J. Phonetics **7**, 279–312 (1979)] and initial consonant VOT [Lisker and Abramson, Lang. Speech **10**, 1–28 (1967)]. A similar effect on initial consonant duration was also found. These effects were stable across speakers, words and contexts, suggesting they may be available as cues to overall word length in online segmentation. To test this, the noun pairs of one speaker were cross spliced to produce congruous and incongruous tokens, and presented to listeners in a "visual world" eye-tracking paradigm [Allopenna *et al.*, J. Mem. Lang. **38**, 419–439 (1998)]. We will report on whether beneficial durational cues of congruous tokens and/or potentially misleading incongruous cues had more influence on online word segmentation of our subjects. [Work supported by NIH and NSF.]

1aSC7. What happens to segment durations at the end of a word?
Rebecca Schwarzlose and Ann R. Bradlow (Dept. of Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208)

Coda consonant clusters of monosyllabic words were studied to determine what durational modifications occur at the segment level as a function of the number, the order, and the morphemic status of segments in the coda. We focused on the segment /s/ because it has the fewest constraints in terms of position in coda clusters, allowing reversals such as tax (/tæ ks/) and task (/tæsk/), and because /s/ can also be a separate morpheme, allowing us to compare segment durations for tax and tacks (/tæ ks/). Preliminary results from four talkers reaffirm previous findings that segment durations decrease as the number of segments increases, although all segments do not decrease to the same extent. Our data also reaffirm previous claims that morphemic /s/ has a longer duration relative to the overall word duration than its nonmorphemic counterpart. In addition to these factors, position at the right-edge boundary of a word seems to play a special role in determining the relative durations of segments within a coda cluster. Specifically, we found that the balance between segment durations in /sk/ and /ks/ coda clusters in monomorphemic words tended to tip in favor of the word-final segment and against the word-internal member of the cluster.

1aSC8. Use of formant movement detail in vowel identification. Amy T. Neel, Ananthanarayan Krishnan, and Kent J. Collins (Dept. of Audiol. & Speech Sci., Purdue Univ., West Lafayette, IN 47906)

Many studies have shown that American English listeners identify vowels with dynamic formants better than flat-formant vowels. The amount of formant movement detail used by listeners, however, is not known. This study examined use of formant movement detail in vowel identification and discrimination using five sets of Klatt-synthesized vowel stimuli modeled after natural vowels. One-point stimuli consisted of flat-formant vowels with formant values taken from the center of the natural vowels. The other four sets consisted of tokens with two, three, five, or eleven points having formant values taken from the natural vowels with the remainder of the values being linearly interpolated. Two-point vowels, with sloping formants, were least well identified by listeners, whereas the most detailed stimuli, the five- and eleven-point vowels, were the best identified tokens. Listeners had difficulty discriminating between the five- and eleven-point representations of vowels. Results indicate that listeners may have an intermediate level of formant movement detail representation — more detailed than a simple slope from onset to offset, but less detailed than natural vowel patterns. In addition to behavioral data, frequency-following responses (FFRs), electrophysiologic measures of neural encoding of formant movement patterns, will be presented. [Work supported by 2001 Research Grant in Speech Science, ASHF/ASA.]

1aSC9. Acoustic and articulatory characteristics of Korean affricates.
Hyunjoo Chung and Anna Marie Schmidt (School of Speech Pathol. and Audiol., Kent State Univ., Kent, OH 44242, hchung@kent.edu)

The purpose of this study was to examine the three voiceless Korean affricates in terms of differing acoustic and articulatory characteristics which might determine three-way lax, aspirated, and tense distinctions. In previous studies, the three-way stop distinction and the two-way fricative distinction (lax and tense) were investigated. Yet little is known about how the three distinctions are implemented in affricates. The three Korean affricates in CV and VCV syllables produced by two female and one male native Korean speaker were compared. Duration, amplitude, and fundamental frequency in six vowel contexts for each affricate were measured on spectrograms. In addition, temporal and spatial characteristics of lingual gestures during production of affricates such as seal duration and skewness of articulatory movement were measured from electropalatographic data. Endoscopic and air flow data for two different subjects were also examined for differences in measures such as glottal width and air flow. Finally, measurements of affricates will be reported and discussed in

comparisons to previous findings on stops and fricatives with respect to the three-way distinctions. [Speech Communication Best Student Paper Award.]

1aSC10. Some acoustic cues for categorizing American English regional dialects. Cynthia G. Clopper and David B. Pisoni (Speech Res. Lab., Psych. Dept., Indiana Univ., Bloomington, IN 47405, cclopper@indiana.edu)

The perception of phonological differences between regional dialects of American English by naive listeners is poorly understood. Using the TIMIT corpus of spoken sentences produced by talkers from a number of distinct dialect regions in the United States, an acoustic analysis conducted in Experiment I confirmed that several phonetic features distinguish between the dialects. In Experiment II recordings of the sentences were played back to naive listeners who were asked to categorize each talker into one of six geographical dialect regions. Results suggested that listeners are able to reliably categorize talkers into three broad dialect clusters, but have more difficulty accurately categorizing talkers into six smaller regions. Correlations between the acoustic measures and both actual dialect affiliation of the talkers and dialect categorization of the talkers by the listeners revealed that the listeners in this study were sensitive to acoustic-phonetic features of the dialects in categorizing the talkers. Taken together, the results of these experiments suggest that naive listeners are aware of phonological differences between dialects and can use these differences to categorize talkers by dialect. Implications for the study of variation and variability in speech will be discussed. [Work supported by NIH.]

1aSC11. Similarities in acoustic modeling for /r/ and /l/. Carol Y. Espy-Wilson (Dept. of Elec. and Computer Eng., Boston Univ., Boston, MA 02115, espy@bu.edu) and Suzanne E. Boyce (Univ. of Cincinnati, Univ. of Cincinnati, Cincinnati, OH 45267)

It is well known that acoustically American English /r/ and /l/ show similar $F1$ and $F2$ values, with /l/ showing a high $F3$ and /r/ showing a low $F3$. Stevens [Acoustic Phonetics, 1998] models tip-down bunched /r/ and light /l/ with similar tongue configurations and lateral acoustic paths around the tongue. Two major articulatory differences in this model are the more forward position of the tongue/palate contact for /l/ vs /r/ and the longer constriction for bunched /r/ vs /l/. Ong and Stone [Phonoscope **1** (1998)] note that the /l/ and /r/ in their data showed very similar pharyngeal volumes, accounting for the similarity in $F1$–$F2$ patterns for these segments. In this paper, we explore the possibility that these variants of /r/ and /l/ share acoustics mechanisms for generating $F1$ and $F2$, and that the difference in $F3$ for /l/ vs /r/ arises primarily from palatal constriction length and front cavity shape.

1aSC12. Acoustic properties underlying vowel intelligibility in words with mixed consonantal contexts. Sarah Hargus Ferguson, Diane Kewley-Port, and Aaron P. Brown (Dept. of Speech & Hearing Sci., Indiana Univ., 200 S. Jordan Ave., Bloomington, IN 47405, safergus@indiana.edu)

This study explored the relative importance of spectral target, dynamic formant, and duration information to vowel identification in mixed consonantal contexts. Vowel intelligibility was varied naturally by one talker producing test words in sentences using clear and conversational speech. In /bVd/ context only, vowels identified by normal-hearing listeners in noise were significantly more intelligible (15%) in clear speech than in conversational speech. Clear and conversational tokens of the 10 vowels also differed significantly on several spectral and temporal measures. Regression analyses revealed that intelligibility differences for individual vowels were well predicted by various combinations of spectral target, dynamic formant, and duration metrics [S. H. Ferguson and D. Kewley-Port, J. Acoust. Soc. Am. **106**, 2272 (1999)]. A second data set was col-

lected concurrently with the /bVd/ data, with the same listeners identifying vowels in mixed CVC contexts. A significant intelligibility advantage (18%) for clear speech was found for these materials. However, regression analyses were less straightforward for vowels in CVC context than in /bVd/ context. Several acoustic metrics were investigated to determine whether intelligibility differences could be modeled for vowels in mixed consonantal contexts. The results will be discussed in terms of theories of vowel perception. [Work supported by NIHDCD-02229, the American Speech-Language-Hearing Foundation, and NIHDCD-00012.]

1aSC13. Influence of /r/ on burst spectra for stop place. Megan M. Hodge, Janka Hegedus, Megan Kenny, and Carmen Souster (Dept. of Speech Pathol. and Audiol., Rm. 2-70 Corbett Hall, Univ. of Alberta, Edmonton, AB T6G 2G4, Canada)

Bilabial and lingua-alveolar stops in 20 minimally contrastive word pairs in both singleton and stop /r/ (cluster) contexts were recorded from two girls, one with facial paralysis (CFP) and one with normal facial movement (CNM). Auditory identification of these productions by 12 listeners revealed a significant place by context interaction for CFP. Identification scores were high for her lingua-alveolar stops (99.3% singletons; 94.8% clusters). Identification scores for her bilabial stops were lower, with singletons being significantly lower (57.8%) than clusters (77.8%). Acoustic cues for stop place ($F2$ onset frequency, VOT, mean and skewness of burst spectra) were measured for all word productions. For both girls, $F2$ onset and VOT measures were lower for bilabial than lingua alveolar stops in singletons and clusters. CFP's burst spectra for bilabials had a higher mean and more negative skewness than lingua alveolars in singletons and clusters. CNM's burst spectra for bilabials had a lower mean and more positive skewness than lingua alveolars in singletons (expected) but a higher mean and more negative skewness for bilabials than lingua alveolars in clusters (unexpected). This unexpected finding may account for CFP's higher identification rates for bilabial clusters (all acoustic cues ''fit'' those of a child with normal facial movement).

1aSC14. A method for studying variability in fricatives using time-dependent changes in spectral mean. Benjamin Munson (Dept. of Commun. Disord., Univ. of Minnesota, 115 Shevlin Hall, Minneapolis, MN 55455)

A large literature has examined variability in temporal, spectral, and kinematic measures as indices of speech motor control. While much of this literature has examined factors that influence within-speaker variability in temporal parameters, relatively little research has examined variability in acoustic parameters, and, in particular, variability in consonant acoustics. This paper presents a new method for studying within-speaker variability in fricative production by examining dynamic characteristics of the spectral mean. In this method, nonlinear regression is used to predict values of the first spectral moment of a 10-ms window of frication noise from its position in the fricative. The resulting measure of model fit, R^2, is used as an index of within-speaker variability. This method was used on two corpora of archival data. In the first study, three English-speaking adults' productions of /s/ in initial /sp/ and /sw/ clusters was examined. /s/ was found to be spectrally less variable when embedded in /sw/ sequences than when embedded in /sp/ sequences. The second study examined variability in the spectra of /f/, /s/ and /ʃ/ spoken at normal and slow rates of speech by adults and two groups of children. Age-group, rate and phonetic context affected fricative variability.

1aSC15. Evaluating vowel normalization transformations through generalized linear modeling. Patti Adank (Dept. of Linguist. and Dialectology, Univ. of Nijmegen, E9.20, Erasmuslaan 1, 6500 HD, Nijmegen, The Netherlands) and Roel Smits (Max Planck Inst. for Psycholinguist., Nijmegen, The Netherlands)

Nearey [ICSLP (1992)] suggests using generalized linear modeling (GLM) to investigate the role of vowel-intrinsic (e.g., role of $F0$) and vowel-extrinsic (e.g., z-transformation, log-mean transformation) normalization schemes. A generalized linear model consists of two parts: a linear predictor and a nonlinear link function. GLM is used to predict values both for dependent variables with discrete distributions and for variables that are nonlinearly related to the predictors. GLM can therefore be useful for dealing with relations between vowel-intrinsic and vowel-extrinsic information and vowel categories, since intrinsic factors can be tested as predictors in the linear predictor function and extrinsic factors can be tested using the link function. The GLM approach was applied to speech data from 80 speakers from four different regions in The Netherlands. The regional accent spoken in each group was different from the other three groups. The nine monophthong vowels of Dutch were produced in a neutral context. For each vowel, $F0$ and $F1-F3$ were measured. These values were transformed according to several vowel normalization schemes. The coefficients for the normalizations were estimated through maximum likelihood estimation. Overall, the results indicate that GLM can be a useful tool for evaluating normalization transformations.

1aSC16. Temporal trading relationship between singletons and geminates in different vowel contexts. Mohamed Al-Khairy (Prog. in Linguist., P.O. Box 115454, Univ. of Florida, Gainesville, FL 32611-5454, mohamedk@ufl.edu)

Length is phonemically contrastive in both consonants and vowels in Arabic. The goal of this study is to examine whether a trading relationship exists between the duration of singletons and geminates in the environment of short and long vowels in Arabic. Preliminary results based on duration measurements of 48 tokens of word-medial bilabial stop singletons and geminates spoken by four speakers suggested that, as expected, geminates are always longer (2.6 times on the average) than singletons. Long vowels are always longer (2.45 on the average) than short vowels. A trading relationship, however, exists between consonant and vowel durations. First, short vowels became significantly shorter when occurring before geminates. This is true only when the vowel following the geminate is also short. Second, geminates preceding short vowels are significantly longer (3.2 vs 2.3 times on the average) than geminates preceding long vowels. To see if these preliminary results also hold true for other types of stop consonants in Arabic, more data will be collected. Similarities and differences in the trading relationship between these consonants and vowels will be examined. Cross-language comparisons regarding the trading relationship between singletons and geminates, and their neighboring vowels will be discussed.

1aSC17. Identification of Swedish vowel quantity by pre-adults. Dawn M. Behne (Norwegian Univ. of Sci. and Technol., NO-7491 Trondheim, Norway), Peter E. Czigler, and Kirk P. H. Sullivan (Umeaa Univ., SE-901 87 Umeaa, Sweden)

Research on the perception of Swedish vowel quantity by adult native listeners shows that vowel duration is a primary cue for identifying vowel quantity. However, when a vowel has a relatively long duration (e.g., due to its inherent duration, or its context), adult listeners may make use of both duration and spectra to identify vowel quantity. Use of the vowel spectrum in these special cases might be seen as the result of perceptual fine tuning to improve the processing efficiency in vowel identification. If so, we would expect a developmental change in how children use vowel duration and spectra to identify vowel quantity, with the use of vowel spectra as a cue coming relatively late. This investigation examines how young native listeners, from 9 to 15 years old, use vowel duration and $F1$-$F2$ to identify Swedish vowel quantities. For each of three vowel types,

100 /kVt/ words were resynthesized with ten degrees of vowel duration and ten degrees of $F1$ and $F2$ adjustment. Each child responded to a subset of these in a rhyming task. Results show a late developing ability to identify vowel quantity and trace the use of vowel duration and $F1$-$F2$ as perceptual cues to vowel quantity.

1aSC18. Quality of front vowels before /r/. Michael J. Clark and James M. Hillenbrand (Speech Pathol. and Audiol., Western Michigan Univ., Kalamazoo, MI 49008)

The vowels /i/ and /ɪ/ are not contrastive before /r/ in American english, and the phonetics literature is equivocal about which symbol to use for the nucleus in words such as *beer*. Similarly /e/ and /ɛ/ are not contrastive before /r/, and the literature contains varied references to one or the other in words such as *bear*. Trained listeners disagree about representing the vowel in /r/-final words. Ten men and ten women recorded monosyllables containing /i,ɪ,e,ɛ/, with initial /b/ and /h/ and with no final consonant or with final /d/ and /t/. For example, the /b/-initial high front series was *bee*, *bead*, *beet*, *bid*, *bit*. Additionally an /r/-final set (*beer*, *bear*, *hear*, *hair*) was recorded. Acoustic measurements included the first two formants. Formants for /i,ɪ,e,ɛ/ in /d/-final and /t/-final syllables (and for /i,e/ in open syllables) were typical of those measured in an earlier study of subjects from the same region. The high front vowel in /r/-final syllables (*beer* and *hear*) showed formant values intermediate between /i/ and /ɪ/, but closer to /i/. The mid front vowel (in *bear* and *hair*) fell entirely within the range seen for /e/. Listening tests using synthetic vowels and backward versions of the original syllables will be discussed. [Work supported by NIH.]

1aSC19. Duration and rate effects on English vowel identification by native Spanish listeners. Terry L. Gottfried and Adam M. Berman (Dept. of Psych., Lawrence Univ., Appleton, WI 54912)

Several American English vowel contrasts are specified by co-varying formant frequency and duration information. Listeners alter their vowel identification according to speaking rate, apparently making judgments about relative vowel duration. Two series of spectral continua were previously created from natural speech, ''beat-bit'' and ''pat-pet,'' varying $F1$ and $F2$, and syllable duration [Berman *et al.*, J. Acoust. Soc. Am. **105**, 1402 (1999)]. These syllables were inserted into sentence contexts of two rates (normal and fast). Longer syllable duration led to more ''beat'' and ''pat'' responses; faster rate contexts also led to more ''beat'' and ''pat'' responses. The present research tested native speakers of Spanish on their use of duration and rate context in English vowel identification. Despite the absence of Spanish vowel contrasts based on duration, native Spanish listeners were significantly affected by syllable duration and sentence rate in their identification of the ''beat-bit'' contrast. They also showed a significant duration effect for ''pat-pet,'' but the magnitude of this duration effect was significantly smaller for these non-native listeners than for native listeners, and there was no significant effect of sentence rate. With less English experience, non-natives might show even smaller effects of duration and sentence rate.

1aSC20. The acoustics of gestural overlap. Khalil Iskarous (Haskins Labs., 270 Crown St., New Haven, CT 06511)

Acoustic modeling of speech production focuses on the effects of a single tongue constriction and a simultaneous labial constriction on the formant frequencies of an acoustic tube. But research on dynamic speech production reveals that tongue gestures overlap in time, so two or more simultaneous tongue constrictions affect the acoustic output of the tube. Understanding the acoustics of natural speech therefore requires an investigation of the acoustic effects of multiple constrictions. The goal is to generate a generalized formant nomogram that would reveal how the acoustic effects of single constrictions combine and interact in determining the acoustic output of the whole tract. Results will be presented from

a finite element simulation of the acoustic effects of two constrictions that vary independently in location and size. The location of each constriction varied in 34 steps and the size varied in 16 steps. Independent variation of the two constrictions yields 295 936 constriction combinations. The function relating the acoustic output of the multiply constricted tract to the acoustic effects of each of the constrictions when made alone is generally nonlinear. Several features of this function will be discussed, and conditions for its linearization will also be presented. [Work supported by NIH.]

1aSC21. Center of gravity effects in the perception of virtual formant transitions. Nandini Iyer, Ewa Jacewicz, Lawrence L. Feth, and Robert A. Fox (Dept. of Speech and Hearing Sci., The Ohio State Univ., 1070 Carmack Rd., Columbus, OH 43210)

When $F2$ and $F3$ of vowels are separated by up to 4.3 Barks (6.3 ERBu) and amplitude modulated to change their spectral centroid over time, listeners perceive a transition similar to that in two-formant diphthonglike vowels [Lublinskaja, 1996]. This study replicates Lublinskaja's finding relating to Russian vowels and extends its application to the consonant-vowel (CV) transitions in /wi/ and /ju/ in English. Stimuli were two-formant or three-formant CVs generated by an upward (/wi/) or downward (/ju/) transition of $F2$ over the frequency range $\Delta f = 100-800$ Hz, in steps of 100 Hz. For the three-formant CVs, $F2$ and $F3$ were separated by the same Δf as in the two-formant CVs, and their amplitudes were modulated so as to move the spectral centroid up or down. The differences in amplitude between $F2$ and $F3$ were either 10 dB or 20 dB. The stimuli were 88 ms in duration. Listeners identified the stimuli in a one-interval, forced-choice identification procedure. Results for the English CV combinations are discussed in relation to the findings reported in Lublinskaja's study. [Research supported by a grant from The Ohio State University College of Social & Behavioral Science to L. Feth and an INRS Award from NIH to R. Fox.]

1aSC22. Perception of relative formant amplitude for /i/ and /ɪ/ in English. Ewa Jacewicz (Dept. of Speech and Hearing Sci., The Ohio State Univ., 110 Pressey Hall, 1070 Carmack Rd., Columbus, OH 43210)

This work investigates listeners' sensitivity to variations in the intensity of $F2$ and $F4$ for the English /i/-/ɪ/ distinction using four-formant synthetic vowels. Two nine-step /i/-/ɪ/ continua were synthesized with the amplitudes of $F2$ and $F4$ increasing or decreasing in steps of 2 dB. During 2AFC discrimination tasks, listeners were able to detect the 2-dB difference for both formants. /i/ was perceived as more natural with decreasing intensity of $F2$ and as less natural with increasing intensity of $F2$. The opposite was observed for /ɪ/. Stepwise intensity changes of $F4$ were perceived as a difference in timbre for both vowels but not as a difference in naturalness. These data suggest that, perceptually, the patterns of relative formant amplitude provide information about the naturalness of /i/ and /ɪ/ in English. The selective auditory sensitivity to the differential contribution of the intensities of $F2$ (variation in naturalness) and $F4$ (variation in timbre) is manifested when the 2-dB difference is detected for both frequency peaks. [Work supported by an INRS Award from NIH to R. Fox.]

1aSC23. Spectral tilt versus formant frequency in static and dynamic vowels. Michael Kiefte and Keith R. Kluender (Dept. of Psych., Univ. of Wisconsin–Madison, Madison, WI 53706)

The relative importance of gross (e.g., spectral tilt) versus detailed (e.g., formant peak frequencies) for speech perception is an enduring question of substantial importance in understanding performance of both impaired and unimpaired listeners [cf. Kiefte, J. Acoust. Soc. Am. (submitted)]. In this study, subjects were presented with steady-state stimuli varying perceptually between [i] and [u] for which spectral tilt and formant frequencies were fully crossed. Results show an effect for both tilt and formants indicating some perceptual importance for both in the iden-

tification of statically specified vowels. However, this result does not appear to extend to more natural sounding stimuli with dynamic formants. For example, there was no effect for spectral tilt in a similar experiment using diphthongs ranging from [ai] to [au]. One hypothesis for this difference is that the F_2 of front vowels is masked by F_1 when combined with sharply falling spectral tilt and that this effect is mitigated in dynamic or naturally produced vowels. Results are also presented which compare the perception of these continua with similar stimuli in which F_2 is absent as well as with dichotically presented stimuli in which F_1 and F_2 are heard in different ears. [Work supported by NIDCD DC04072.]

1aSC24. Vowel formant discrimination in quiet and noise for natural speech. Chang Liu and Diane Kewley-Port (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405)

Everyday communication takes place in background noise. The purpose of this study is to investigate the effects of noise on formant discrimination for naturally spoken vowels. Our analysis of formant discrimination for synthetic vowels found that formant resolution could be accounted for by excitation pattern and specific loudness models [D. Kewley-Port and Y. Zheng, J. Acoust. Soc. Am. **103**, 1654 (1998)]. It is expected that excitation pattern and specific loudness model can be extended to account for the formant discrimination in natural speech in both quiet and noise, even though natural speech has more variability than synthetic speech. The present study measured formant frequency discrimination in quiet and noise for normal-hearing listeners. Thresholds were measured as ΔF in natural speech for $F1$ and $F2$ for a set of vowels /ɪ, ɛ, ʌ, æ/ spoken by a female talker. Noise conditions used a long-term averaged speech spectrum to shape the noise. The signal-to-noise ratio (SNR) was manipulated systematically from -6 to $+6$ dB. Results showed different masking

patterns for $F1$ and $F2$ with performance approaching asymptote between 0 and $+3$ dB SNR. Results from an analysis of these data with a specific loudness pattern model will be presented. [Work supported by NIHDCD-02229.]

1aSC25. Dynamic information in the perception and production of Rhode Island vowels. Harriet Magen (Haskins Labs, 270 Crown St., New Haven, CT 06511 and Dept. of Commun., Rhode Island College, Providence, RI 02908, hmagen@ric.edu) and Khalil Iskarous (Haskins Labs, New Haven, CT 06511)

Comparisons were made of the production and perception of 11 vowels for 10 Rhode Island speakers. To assess perception, the method of adjustment (MOA) was used. Subjects were presented with a grid of 298 steady state $F1/F2$ combinations and asked to choose the best match to a visually presented keyword. [Johnson *et al.*, Language **69**, 505–528 (1993)]. The same subjects produced the 11 vowels; acoustic measurements were made 1/3 of the way through. For two pairs of vowels ([eɪ]-[ɪ], [ɔ]-[oʊ]) there was a disparity between production and perception, calling into question the perceptual strategy listeners use when normally present dynamic information is absent. To address the issue of the dynamics of production, measurements were made at additional points. At all three points measured, values of $F1$ of [eɪ] and [ɪ] overlap substantially; however, perceptually they are distinct. While this dialect maintains the [ɑ]-[ɔ] distinction, [ɔ] and [oʊ] show overlap in production early in the vowel, diverging later in the vowel; static formant values chosen as the best instance of [oʊ] overlapped with the formants of producted [ɔ]. The efficacy of the MOA for dynamic information will be discussed. [Work supported by NIH and the RIC Faculty Research Fund.]

MONDAY MORNING, 4 JUNE 2001

PDR 18, 7:45 TO 10:50 A.M.

Session 1aSPa

Signal Processing in Acoustics: Signal Processing for Underwater Acoustics

Brian G. Ferguson, Chair

Maritime Systems Division, Defence Science Technology Organization, P.O. Box 44, Pyrmont 2009, Australia

Chair's Introduction—7:45

Contributed Papers

7:50

1aSPa1. Statistical analysis of acoustic Green's functions in a shallow ocean environment. Sean R. Chapin, George E. Ioup, Juliette W. Ioup (Dept. of Phys., Univ. of New Orleans, New Orleans, LA 70148, schapin@uno.edu), and George B. Smith (Naval Res. Lab., Stennis Space Center, MS 39529)

Knowledge of statistical properties of acoustic Green's functions can assist in the development of effective blind deconvolution algorithms for enhancing the detection and classification performance of active and passive sonar systems. One hundred synthetically generated Green's functions have been produced representing the acoustic propagation of signals in a given shallow ocean environment. Fourier analysis is performed on the Green's functions and various statistical properties are calculated in both the time and frequency domains, including examination of properties such as the mean, variance, and the third- and fourth-order moments in both domains. Statistics are also generated for the entire ensemble of Green's functions as a function of time and frequency. A similar statistical analysis is conducted on the autocorrelation of each of the Green's func-

tions in both domains. Synthetic noise-free signals are produced by convolving the Green's functions with representative source signals. Statistical analysis is performed on the resulting signals in both domains for comparison to the results obtained using the Green's functions alone. Results include large phone-to-phone variations in kurtosis of the Green's functions and ensemble average kurtosis increasing sharply with time and decaying to zero. [Research supported by ONR.]

8:05

1aSPa2. Improved detection using propagation and target responses. Charles F. Gaumond and Ralph N. Baer (Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375-5320)

The detection results of a monostatic sonar simulation are presented. The bandwidth of the simulation is 200 Hz to 950 Hz and the environment is range-independent, shallow water. A target is modeled as a point with a frequency-dependent response simulated using a finite element model and the propagation is simulated using the RAM PE model. These two broad-

band simulation results are combined using a coherent form of the sonar equation. The improvement of detection is shown for several cases comparing the use of environment-alone, target-alone, and combined-environment-target matched filters with the detection from the use of simple matched filtering. [Research sponsored by ONR.]

8:20

1aSPa3. Wavelet detection and denoising of low-frequency chirp signals. Joseph S. Wheatley, Juliette W. Ioup, and George E. Ioup (Dept. of Phys., Univ. of New Orleans, New Orleans, LA 70148, wheatley@datastar.net)

The detection and classification of underwater acoustic signals embedded in noise is a fundamental problem of interest to the Navy. The use of wavelet transforms is a recent development in digital signal processing that has been applied in many different areas. A particular type of wavelet is the chirplet, which includes frequency variation as well as time shift and scaling. The analysis of low-frequency signals containing multiple chirps using wavelet and chirplet techniques is demonstrated. Examples of low-frequency synthetic chirp signals are generated. Denoising and feature extraction of these signals using various wavelets with wavelet packet techniques are described.

8:35

1aSPa4. Time-reversal operator applied to broadband active sonar. David M. Fromm and Charles F. Gaumond (Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375-5320)

The decomposition of the time reversal operator (DORT) is a single-frequency technique that isolates scatterers using a system of multiple sources and receivers [Prada *et al.*, J. Acoust. Soc. Am. **99**, 2067–2076 (1996)]. The application of this technique to a broadband active sonar system is shown through simulations in a realistic oceanic environment using RAM PE and a coherent reverberation model. A broadband signal containing reverberation, noise, and a target return is simulated. The broad bandwidth allows point targets or clutter to be isolated in range. The returns of interest are then time windowed, Fourier transformed, and DORT processing is applied. By propagating the eigenvectors of the time reversal operator to create time reversal images, the character of the return can be evaluated. [Research sponsored by ONR.]

8:50

1aSPa5. Iterative deconvolution for source recovery from multipath degraded signals. James P. Larue, George E. Ioup, Juliette W. Ioup (Dept. of Phys., Univ. of New Orleans, New Orleans, LA 70148, hubauto@juno.com), and George B. Smith (Naval Res. Lab., Stennis Space Center, MS 39529)

Known ocean propagation Green's functions can be used to recover an estimate of the source in underwater acoustic passive detection. Good estimates of the source improve the performance of classification algorithms. Since convolution models the propagation, deconvolution using the Green's function as the impulse response produces an estimate of the source. In the presence of moderate amounts of noise, deconvolution methods degrade due to the extreme noise sensitivity of the process. Iterative deconvolution has been shown to be more robust in the presence of noise than some other techniques due to the effect of a gradually changing window that need not be pushed to convergence. Iterative techniques are more difficult to apply for oscillatory data than for non-negative data

because conservation of area does not provide a robust normalization. Iterative deconvolution techniques are tested for noise performance in source recovery and compared to other methods. [Research supported by ONR.]

9:05

1aSPa6. Matched field processing classification. Ahmad T. Abawi, Phil Schey (53560 Hull St., San Diego, CA 92152-5001), and W. S. Hodgkiss (Scripps Inst. of Oceanogr., San Diego, CA 92152-6400)

Matched field processing (MFP) has traditionally been performed using vertical line arrays. During the Shallow Water Cell Experiment SWellEx-96, which was conducted off the coast of San Diego in the spring of 1996, it was shown that MFP could also be done using horizontal line arrays (HLA). The MFP correlation using the 250-m long HLA used in this experiment was large enough for detection, but because of the high sidelobes, classification by depth was not possible. However, simulations predicted that the sidelobe levels for an approximately 500-m long HLA would be small enough to make classification by depth possible. To verify this prediction, the SWellEx-99 experiment was conducted in spring of 1999. During this experiment, which was conducted in approximately 200 m of water, a number of large ($\sim$500 m) horizontal line arrays were deployed. A J15-3 source, which transmitted multi-tone signals at various levels, was towed along various tracks. In this paper, data from the SWellEx-99 experiment are used to examine the detection and classification performance of adaptive matched field processing on large horizontal line arrays.

9:20

1aSPa7. Robust adaptive beamforming for hydrodynamic self-noise rejection. Vincent E. Premus and Stephen M. Kogon (MIT Lincoln Lab., Lexington, MA 02420)

White noise gain (WNG) is a metric widely used in the radar community for limiting the degree of adaptivity of an adaptive beamformer (ABF). Constraining adaptation of an ABF algorithm has two key effects: (1) protection against self-nulling associated with steering vector mismatch, and (2) limitation of white noise gain associated with squinting a beampattern to place a null on a mainlobe interferer. Hydrodynamic self-noise, or cable strum, commonly manifests itself as a source of mainlobe interference for passive acoustic towed horizontal line arrays. Cable strum is the result of vibrations excited in the array body by vortex shedding in the presence of flow. Strum is particularly detrimental on beams near endfire and typically exhibits very high dynamic range. As such, it requires an aggressive adaptation strategy for its effective removal. In this work, a white noise gain constraint for the rejection of cable strum is derived based on the scaled orthogonal projection approach of Cox *et al.* [IEEE Trans. ASSP **35** (1987)]. Tradeoffs involved in balancing strum rejection performance against mismatch-induced signal self-nulling are examined. [Sponsored by the Dept. of the Navy, under Air Force Contract No. F19628-00-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and not necessarily endorsed by the U.S. Air Force.]

9:35

1aSPa8. Effective analysis technique of unstable acoustic signature from ship radiated noise. Jong-rak Yoon, Yong-ju Ro (Telematics Eng. Dept., Pukyong Natl. Univ., Pusan 608-737, Korea), and Jae-Jin Chun (Agency for Defense Development, Chinhae 645-600, Korea)

Acoustic signatures of ship radiated noise are important passive sonar parameters in target ship detection and classification. Among these signatures an unstable signature, which is defined as frequency change with respect to the time or frequency modulation, is due to the external loading

variation in specific machinery components. As an example, propeller blade frequency is measured as an unstable signature and is controlled by the speed governor which is dependent on the propeller resistance to sea surface waves or current turbulence. In this study an extended Kalman filtering algorithm is applied to detect the unstable signature. To obtain better performance than with the existing similar technique, state model parameters such as variance and bandwidth are expressed as a function of the sea state wave spectrum. In addition background noise variance in the measurement model is also defined as a sea state function. Its performance is evaluated using measured surface ship radiated noise data and it is also applied to Dopplerized signal analysis which is essential in moving target monitoring. [Work supported by the Underwater Acoustics Research Center.]

tion and equalization. In this paper a single-input multi-output (SIMO) channel blind equalizer for an underwater acoustic communication system is considered. A direct blind equalization scheme was investigated by simulations, which is a novel unsupervised adaptive equalizer and was suggested by J. Labat first for single-input single-output (SISO). An interactive matched processing (MP) is added for the SIMO case. The work verified the prediction that the linear equation form of the direct blind equalizer lends itself naturally to adaptive solutions and optimum combinations of the equalizer outputs and improved performance over existing algorithms in terms of mean-square error and probability of error in the equalized data.

9:50

1aSPa9. Wavelet analysis of acoustic Green's functions in a shallow ocean environment. Antonio M. Morgan, Juliette W. Ioup, George E. Ioup (Dept. of Phys., Univ. of New Orleans, New Orleans, LA 70148, ammorgan2@hotmail.com), and George B. Smith (Naval Res. Lab., Stennis Space Center, MS 39529)

Knowledge of the properties of acoustic Green's functions can assist in the development of effective blind deconvolution algorithms for enhancing the classification performance of active and passive sonar systems. Green's functions have been produced representing the acoustic propagation of signals in a given shallow ocean multipath environment. Simulated noise-free received signals are the result of convolving the Green's functions with representative source signals. Wavelet analysis is performed on the Green's functions, the source signals, and the received signals. Forty different wavelets are tested. Comparison of the wavelet transform coefficients illustrates similarities and differences in signal properties. Significant differences can be exploited to separate the source from the Green's functions in the multichannel received signals for blind deconvolution source recovery. [Work supported by ONR.]

10:05

1aSPa10. An experimental investigation on a direct blind equalizer for underwater acoustic multiple FIR channel. Jianmin Liu, Yingmin Wang, Xiangrui Cao, and Zhengqi Zhang (Inst. of Acoust. Eng., Northwestern Polytech. Univ., Xi'an 710072, PROC)

A blind equalizer removes the distortion caused by intersymbol interference (ISI) without the need for training sequences. In order to get the complete phase information of the ISI, high-order statistics have to be used to find the equalizer taps. However, L. Tong *et al.* have shown that most channels can be identified from the second-order cyclostationary statistics through fractionally spaced equalizers. This result has stimulated considerable research activities in fractionally spaced channel identifica-

10:20

1aSPa11. Color rendering of range in passive acoustic displays. Frank A. Boyle and Thomas H. Phipps (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029)

Techniques for mapping image color to range in passive time-bearing displays were explored. They included (1) a range-focusing method wherein images focused at multiple ranges were combined; and (2) a parallax technique wherein images were formed with different subarray apertures. In both cases, a Hough transform based image processing method was used to enhance features and to compensate for degradation associated with combination of multiple images. Colormapping techniques were employed to render source ranges in color. The presentation will include a description of the methods and a discussion of their effectiveness under various conditions. [Work funded under ARL:UT IR&D program.]

10:35

1aSPa12. Characterization of the left/right bearing ambiguities for passive line arrays. Frank A. Boyle and David E. Grant (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029)

The symmetry of a straight horizontal line array produces a left/right bearing ambiguity that makes it impossible to determine on which side of the array a source is located. A simple technique for resolving left/right ambiguities is to break the symmetry by curving the array. Under certain common conditions, however, significantly curved arrays still possess left/right ambiguities that depend on the source's range and the beamformer's focal length. These ambiguities can be modeled, given a source and array configuration, to produce a prediction of source localization capability. Comparisons with actual data were generally consistent with modeled results in the cases considered. The presentation will include a description of how the ambiguities are modeled, and a comparison between modeled and experimental results. Considerations on how curved arrays can be configured to minimize ambiguities will also be discussed.

Session 1aSPb

Signal Processing in Acoustics: Signal Processing for Airborne Acoustics

Charles F. Gaumond, Chair

Naval Research Laboratory, Code 7142, 4555 Overlook Avenue, SW, Washington, DC 20375-5320

Chair's Introduction—11:00

Contributed Papers

11:05

1aSPb1. Signal processing for acoustic surveillance of the land environment. Brian Ferguson (Defence Sci. and Technol. Organisation, P.O. Box 44, Pyrmont NSW 2009, Australia, Brian.Ferguson@dsto.defence.gov.au)

Surveillance systems that rely on the propagation of sound in the atmosphere are not constrained by electromagnetic line of sight and offer the advantages of being passive, lightweight, compact and covert. Such systems are unattended in their operation and are often deployed in remote areas. To demonstrate the scientific principles that underpin the operation of such systems, various narrow-band and broadband signal processing algorithms are applied to real acoustic data from both single and multiple sensors. In one example, the acoustical Doppler effect enables the flight parameters and blade-passage frequencies of turboprop and rotary wing aircraft to be estimated from the time-frequency signal analysis of spectrogram data. In another example, the acoustical Lloyd's mirror effect is used to extract the flight parameters of a jet aircraft. Other examples deal with the application of generalized cross-correlation processing to the detection, localization and tracking of aircraft, ground vehicles and weapon fire. Wideband cross-correlation with differential Doppler compensation is required for the acoustic tracking of high-speed sources when the sensors are widely separated and phase tranform prefiltering is required to suppress the ambiguouos peaks in cross-correlograms caused by the presence of strong narrow-band lines in the source spectra of air and ground vehicles.

11:20

1aSPb2. Finding the right cross-correlation peak for locating sounds in multipath environments with a fourth-moment function. John L. Spiesberger (Dept. of Earth and Environ. Sci., Univ. of Pennsylvania, 240 S. 33rd St., Philadelphia, PA 19104-6316, johnsr@sas.upenn.edu)

To locate calling animals in reverberant environments from recordings on widely separated receivers, a fourth moment "augmented-template correlation function" (ATCF) helps identify which of many peaks in each cross-correlation function is that corresponding to the difference in travel times for the first arrivals (reference-lag). This peak may not be the largest. The ATCF, by providing an approximate correlation between auto- and cross-correlation functions, can be orders of magnitude more efficient in selecting the reference-lag than the alternative of randomly selecting peaks. The ATCF's efficacy increases with the number of paths and their signal-to-noise ratios. [Work supported by ONR.]

11:35

1aSPb3. A reconstruction technique for passive acoustic imaging through a turbulent atmosphere. Jonathan W. Benson, George W. Swenson, Jr. (Wave Propagation Lab., Dept. of Elec. and Computer Eng., Univ. of Illinois at Urbana-Champaign, Champaign, IL 61820), and Yoram Bresler (Coordinated Sci. Lab., Univ. of Illinois at Urbana-Champaign, Champaign, IL 61820)

Interferometric passive acoustic imaging through the near-ground atmosphere is hampered by the effects of turbulence. Traditional techniques aimed at removing the effects of the atmosphere, such as self-calibration and phase-closure, cannot be applied in this case where the distortion is a function of the source coordinates. A method was developed to remove the distortion due to the atmosphere within the context of interferometric imaging. A discrete complex-valued distorting screen was used to model the atmosphere. This type of model is routinely used to model propagation through turbulent media and in this case was found to adequately model the distortion effects seen in field data.Reconstruction is accomplished by joint estimation of the source intensity distribution along with the parameters of the distorting screen. Multiple short integration period images or snapshots are used. The parameters of the screen are allowed to change from one snapshot to another, while the source distribution is assumed constant. The Cramer–Rao bounds of the problem were computed and found to be very favorable. The estimation problem was solved using both maximum likelihood and least-squares methods. Near-perfect reconstruction was obtained for simulated data, and significant image quality gains were obtained with experimental data.

11:50

1aSPb4. Performance bounds of atmospheric acoustic arrays operating in a turbulent medium. Sandra L. Collier and D. Keith Wilson (U.S. Army Res. Lab., AMSRL-CI-EP, 2800 Powder Mill Rd., Adelphi, MD 20783-1197)

The performance bounds of an atmospheric acoustic array operating in a turbulent medium with fluctuations described by von Kármán's spectrum are investigated. This treatment considers a single monochromatic source and no multiple-path effects. The Cramer–Rao lower bounds of the azimuthal and zenith angles of arrival (AOAs) are calculated. The normalized range (to wavelength), turbulence parameters, and signal-to-noise ratio are introduced as additional unknowns. It is found that the Cramer–Rao lower bounds of the AOAs increase significantly in strong turbulence at large values of the normalized range. For weak turbulence and moderate values of the normalized range, the signal-to-noise ratio is the dominating factor. For an incident plane wave, the estimates of the AOAs will decouple from the estimates of the other parameters with the appropriate choice of array geometry; however, the estimates of the azimuth and zenith are coupled to one another due to the constraints of the turbulence model.

1a MON. AM

Session 1aUW

Underwater Acoustics: Scattering

William M. Carey, Chair

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

Contributed Papers

8:15

1aUW1. Issues of analytical continuation of the field interior to the target surface and a more general T-matrix for submerged fluid targets. M. F. Werby (NRL Code 7180, Stennis Space Center, MS 39529 and Dept. of Phys., Catholic Univ. of America, Washington, DC 20064)

The T-matrix or extended boundary condition method developed by P. Waterman for wave scattering from classical targets and extended by a host of researchers is a marvelous generalization of normal mode solutions allowed for separable geometries to extend to nonseparable geometries by means of mode coupling. The method is based on Huygens' principle and partial wave expansions of physical quantities. In general the expansions don't obey orthogonality and one must take care in the way expansions are made. One must never take expansions of a quantity and then represent expansions of their differential forms in terms of the partial wave differential forms using the same expansion coefficients lest one violate a mathematical principle (the Rayleigh hypothesis). One may circumvent any issues without resort to analytical continuation of interior solutions by including all solutions subject to interior and exterior Greens functions. This leads to matrix relations (constraining matrices) between the expansion terms which when used lead to new forms of the fluid T-matrix, for example. The new form differs from Waterman's T-matrix for the case when the constraining matrices are not symmetric. Otherwise the original expression by Waterman is valid. These issues are discussed and examples are presented.

8:30

1aUW2. Backscattering from empty and water-filled tilted cylindrical shells due to leaky helical flexural waves: Comparison and ray theory. Florian J. Blonigen and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814)

For tilt angles smaller than the meridional ray coupling condition previously investigated [S. F. Morse *et al.*, J. Acoust. Soc. Am. **103**, 785–794 (1998)], helical rays on empty and water-filled steel shells can significantly enhance the backscattering. These contributions are compared and modeled in the present work. Such contributions in the water-filled case are found to be weaker in amplitude, which is to be expected since additional energy is lost through radiation leaked to the inner fluid. A modified ray theory taking into account the increased radiation damping was compared with experimental results for the earliest helical wave arrival, yielding satisfactory agreement. However, the energy lost to the interior of the shell may be at least partially recovered, as internal rays can excite additional leaky waves on the shell. Contributions from rays with one internal chord inside the cylinder were superposed with those of the first helical wave, and the resulting backscattering amplitude was nearly that of the air-filled case for the first helical wave alone. This shows that we may regard the ray theory for the empty shell as an upper bound for corresponding amplitudes observed in the water-filled case. [Work supported by ONR.]

8:45

1aUW3. Wavefront scatter and wavefront healing. C. Feuillade (Naval Res. Lab., Stennis Space Center, MS 39529-5004, cf@nrlssc.navy.mil)

While acoustic scattering is essentially a time domain phenomenon, studies of scattering from objects in water, and from boundaries, have often concentrated on time-independent/cw descriptions. There are, however, critical aspects of the scattering problem which become evident only in the time domain. In this paper, time-domain visualizations of impulse scattering from rigid, pressure-release, fluid, and elastic spheres are presented. The results provide important insights into the physical processes which lead to the construction of scattered wavefronts. Using this method, it is possible to time-separate and identify different scattering mechanisms: geometrical scatter; diffraction (which, for rigid spheres, gives rise to the circumferential wave); refraction; and wave conversion in the body of the scattering obstacle. Of particular interest is the characterization of the forward scatter region, where the process of wavefront healing can be directly observed. The presentation will be illustrated by movie animations. [Work supported by ONR.]

9:00

1aUW4. Shape sensitivity calculations for exterior acoustics problems. Gonzalo R. Feijoo and Peter M. Pinsky (Div. of Mech. and Computation, Stanford Univ., Stanford, CA 94305-4040, grfeij@stanford.edu)

We present a method to calculate the derivative of a functional determined by the shape of an object. This functional depends on the solution of a linear acoustic problem posed in an unbounded domain. We rewrote this problem in terms of another one posed in a bounded domain using the Dirichlet-to-Neumann (DtN) map, or the modified DtN map. Using a classical method in shape sensitivity analysis, called the adjoint method, we are able to calculate the derivative of the functional using the solution of an auxiliary problem. This method is particularly efficient because the cost of calculating the derivatives is independent of the number of parameters used to approximate the shape of the domain. The resulting variational problems are discretized using the finite-element method and solved using an efficient Krylov-subspace iterative scheme. Numerical examples that illustrate the efficacy of our approach are presented.

9:15

1aUW5. A two-way coupled mode method, matrix methods, and bandedness of the integration with free scattering solutions and the solution of imbedded objects in a waveguide. M. F. Werby (NRL Code 7180, Stennis Space Center, MS 39529 and Dept. of Phys., Catholic Univ. of America, Washington, DC 20064)

Since the seminal work of Richard Evans in 1983 when he proved to the world the viability of a two-wave general waveguide solution in coupled mode form his method and implementation has been the gold standard of other range-dependent approximations. The method by its very nature is tedious to implement and requires much time in computation.

2299   J. Acoust. Soc. Am., Vol. 109, No. 5, Pt. 2, May 2001

141st Meeting: Acoustical Society of America   **2299**

Often a one-way solution is adequate allowing many to circumvent this more complicated solution. A one-way coupled mode method has been developed at NRL that takes advantage of representation theory in terms of a basis set that leads to analytical forms suitable for use in representing incident fields as input for scattering problems. It is apparent that in many cases the addition of an inclusion renders coupling in both directions in a waveguide and thus the one-way approach is extended to a two-way coupled mode solution and by extension to a generalization of the integrated solution of an object in a waveguide for more general inclusions.

9:30–9:45   Break

9:45

1aUW6. Scattering measurements and interface model validation for a rocky bottom at 2–4 kHz. Raymond J. Soukup (Naval Res. Lab., Code 7144, 4555 Overlook Ave. SW, Washington, DC 20375)

Validation of interface scattering models for the ocean bottom, such as the small slope approximation model of Gragg and Wurmser [J. Acoust. Soc. Am. **109** (2001)] requires acoustic scattering data covering a wide range of grazing angles with minimal sub-bottom penetration. Such data measurements were performed in the frequency band 2–4 kHz over an exposed limestone bottom off the Carolina coast (120 to 310 m water depths) during the Littoral Warfare Advanced Development Focused Technology Experiment 96-2 (LWAD FTE 96-2). Direct-path bottom scattering strengths were obtained at grazing angles ranging from 8 to 75 degrees using data fusion from multiple experimental geometries coupled with careful signal processing. This data processing includes corrections for the surface-reflected and other hybrid paths, and the rapid reverberation decay observed over the length of the pulses at higher grazing angles. The resulting frequency and grazing-angle dependencies exhibit trends consistent with those predicted by the elastic small-slope approximation model of Gragg and Wurmser. [Work supported by ONR.]

10:00

1aUW7. Theory and computation for small-slope scattering from rough elastic ocean floors. Robert F. Gragg and Daniel Wurmser (Naval Res. Lab., Code 7144, Washington, DC 20375-5350)

A bistatic theory is presented for the scattering strength of a random rough interface bounding a uniform medium (either fluid or elastic solid) for cases where the Bragg scale lies within the power-law tail of the roughness spectrum. The physical basis is an inherently reciprocity-preserving, local small-slope theory. A robust numerical implementation is also presented that functions smoothly throughout a wide range of spectral exponents (from 2.4–3.9, at least). The practical result is a significantly improved formulation for the interface component of sea floor scattering. Results are compared to those of perturbation theory, and illustrate the dependence on frequency, environmental parameters, and bistatic geometry. Validation of the model using at-sea data is addressed by Soukup [R. J. Soukup, J. Acoust. Soc. Am. (submitted)]. [Work supported by ONR.]

10:15

1aUW8. A new strategy for applying the parabolic equation to a penetrable rough surface. Daniel Wurmser (Naval Res. Lab., Code 7144, Washington, DC 20375-5350)

The range-dependent parabolic equation (PE) propagates an auxiliary field that roughly corresponds to the square root of the downrange flux. At the ocean bottom, the interface is broken into stair steps, and this auxiliary field is conserved at the vertical interfaces. Because of its success in matching benchmark solutions, this method is widely used to model standard propagation problems in underwater acoustics. However, when the interface is rough at the wavelength scale, the physics of the problem changes, and the standard approach ceases to be practicable. This talk presents a new approach specifically designed to handle this important problem. The Foldy–Wouthuysen transformation generates the PE for this auxiliary field, complete with a full suite of boundary conditions. In the discrete problem, the boundary conditions are used to evaluate the Hamiltonian on the interface, and the result is used to perform the next downrange step. At lowest order, the new theory streamlines the stair step method, and frees it from *ad hoc* impositions such as those spawned by the density jump at the ocean bottom. The approach also systematically provides higher-order corrections. [Work supported by ONR.]

10:30

1aUW9. Effect of bottom ripple on detection of a buried object using shallow-grazing-angle underwater sonar. Raymond Lim (Coastal Systems Station, Code R22, 6703 W. Hwy. 98, Panama City, FL 32407)

Sound penetration into the seafloor at shallow (subcritical) sonar grazing angles has been a topic of much discussion lately because measured levels can be much higher than expected based on simple propagation models assuming a flat seafloor. A popular mechanism for explaining this penetration is based on diffraction of energy into the bottom by seafloor roughness. For objects buried in the bottom, this holds the promise for enhanced long range detection with sonar as long as the backscatter by the roughness itself does not overpower the echo from the object. Here the detectability of a sphere buried under roughness in the form of anisotropic wave-induced ripple is investigated by modifying a scattering solution that combines a stochastic perturbative description of the roughness with a deterministic description of the sphere response. [See Lim *et al.*, J. Acoust. Soc. Am. **107**, 1246–1262 (2000).] Comparison with predictions for an isotropic, power-law seafloor roughness at a given rms amplitude demonstrates greater high frequency detection capability with rippled seafloors over a wide range of illumination angles relative to the ripple direction. Comparisons with experimental data are also discussed. [Work supported by ONR.]

10:45

1aUW10. Scattering from stratified sediments. Anatoliy N. Ivakin (On leave from Andreev Acoust. Inst., Svernika 4, Moscow 117036, Russia, anivakin@yahoo.com)

There is a special class of volume perturbations, which exist within stratified media and can be described in terms of displacement of isosurfaces of these parameters. Scattering from marine sediments with such perturbations is considered using a unified approach to volume and roughness scattering [A. Ivakin, J. Acoust. Soc. Am. **103**, 827 (1998)]. Discrete stratification (layering) is treated as a particular case of a continuous one, where interfaces between layers are replaced by additional (transition) layers with acoustical parameters smoothly dependent on the depth. In the unperturbed medium, the transition layers are flat, which gives a zeroth-order solution. In the first-order approximation, volume perturbations of the aforementioned class are proportional to vertical gradients of corresponding acoustical parameters in the unperturbed medium and to displacement (roughness) of the isosurfaces of these parameters within transition layers. Note that there are only smooth continuous volume inhomogeneities in these layers, that is, there is no real roughness scattering, which is controlled by discontinuity of the acoustical parameters at the interfaces. A new general model is presented, which gives the volume scattering strength as a function of the wave thickness of the transition layers. The model does not require the vertical correlation sclae of inhomogeneities to be small in comparison with the layer thickness. In the case of low frequencies, where the transition layers are thin in comparison with the wavelength, the model provides an exact transition to existing results

for scattering from rough interfaces and, thus, naturally links volume and roughness mechanisms of seabed scattering at low and high frequencies within a unified approach. [Work supported by ONR.]

facility details the frequency-dependent scattering from well-characterized bubbly cylinders, and a good match to data is obtained using an effective medium model for the uniform compressible cylinder. Accompanying scattering data from two types of bubbly spheres (voided polyurethane, gel-filled) are also presented. [Work supported by ONR.]

11:00

1aUW11. The contributors of sound scattering in stratified flows. Victor Prokhorov (Inst. for Problems in Mech. of the RAS, 101(1) prospekt Vernadskogo, Moscow 117526, Russia)

It is generally assumed that marine biomass dominates the backscattering everywhere in the ocean. However, stratified flows form both aggregated and solitary scatterers whose scattering levels are close to those of biologicals. The present work deals with theoretical and laboratory study of scattering phenomena in stratified flows of different types and regimes. Microscale clusters, flat and curved discontinuities, and solitary vortex rings were modeled and examined with both monostatic and bistatic acoustic schemes. Simultaneously they were optically visualized and photographed to provide their direct identification and computation of spatial spectra, and to extract characteristic scales being responsible for scattering. The average scattering levels converted to dimensionless backscattering cross section are of the same order or above those obtained in field measurements. Theoretical study seeks to reveal the mechanisms producing acoustic contrasts and to examine temporal characteristics of scattering. The scattering dynamics essentially depends on whether salinity or temperature stratification dominates the condition under which the scattering inhomogeneities arise. Also, attached internal waves contribution to produce the scattering interfaces was theoretically and experimentally confirmed. Shadowgraphs of scattering structures followed by synchronous scattering profiles, spatial and temporal characteristics of scattering, data of theoretical modeling are presented.

11:30

1aUW13. Measurements of the acoustic properties of bubbly liquids and partially voided polyurethane solids in the resonance regime. Preston S. Wilson, Ronald A. Roy, and William M. Carey (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215)

Complete understanding of sea-surface scattering, shallow water volume scattering, and bottom interaction requires the understanding of frequency-dependent acoustic propagation in bubbly media. Because of the high attenuation encountered at and near the range of resonance frequencies associated with bubble size distribution, experimental measurements in this regime are difficult, and for many specific cases, have not been obtained. Using an impedance tube method similar to one previously described [J. Acoust. Soc. Am. **106**, 2291(A) (1999)], measurements of the acoustic impedance of bubbly fluids in the vicinity of individual bubble resonance have been obtained. The measurement procedure will be briefly summarized, and contrasted with the previously reported method. Results obtained over a range of frequencies and void fractions with known size distributions will be presented and compared to existing theory. In addition, the acoustic behavior of partially voided polyurethane spheres and plates has been studied using the present impedance tube method. Evidence for collective oscillation of these structures will be shown. [Work supported by ONR.]

11:15

1aUW12. Acoustic scattering from bubbly cylinders and spheres. Preston S. Wilson, Ryan D. McCormick, Ronald A. Roy, and William M. Carey (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215)

This presentation deals with acoustic scattering from bubble clouds. In the most general manifestation, the cloud has arbitrary shape and extent and consists of a population of bubbles with arbitrary size distribution; the ensuing scattering problem is considerable. In the asymptotic limit of an acoustically compact cloud consisting of acoustically small bubbles (both conditions are satisfied at sufficiently low frequencies) the situation condenses down to the much simpler problem of scattering from a uniform compressible object, and to lowest order in the Helmholtz number (ka) this object is a sphere. Experimental evidence supporting this approximation was obtained at Lake Seneca [Roy *et al.*, J. Acoust. Soc. Am. **92**, 2993 (1992)], as well as numerous laboratory noise and scattering experiments that followed. We present additional data that supports this view, and extends it to higher order in ka for the case of scattering from a cylindrical bubble cloud. Data obtained at the NUWC underwater test

11:45

1aUW14. Linear acoustics in bubbly liquids from an effective medium theory. Steven G. Kargl (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Proposed corrections to the lowest-order approximation of an effective wave number obtained from Foldy's exact multiple scattering theory [Foldy, Phys. Rev. **67**, 107 (1945)] has sparked renewed interest in linear wave propagation through bubbly liquids. An alternative approach to wave propagation through a bubbly liquid reduces the governing equations for a two-phase medium to an effective medium. Commander and Prosperetti [J. Acoust. Soc. Am. **85**, 732 (1989)], based on this method, derive an expression for the lowest-order approximation to an effective wave number. At this level of approximation the bubbles interact with the mean acoustic field without higher-order rescattering. That is, the field scattered from a bubble may interact with one or more new bubbles in the distribution, but a portion of that scattered field may not be scattered back to any previous bubble. Simple modifications to the results of Commander and Prosperetti lead to a new expression for the effective wave number, which properly accounts for all higher orders of multiple scattering. [Work supported by ONR, Code 321OA.]

Session 1pAA

Architectural Acoustics: The First 80 Milliseconds in Auditoria

U. Peter Svensson, Cochair

Department of Telecommunications, Norwegian University of Science and Technology, Trondheim N-7491, Norway

Jerald R. Hyde, Cochair
Box 55, Helena, California 94574

Chair's Introduction—1:00

Invited Papers

1:05

1pAA1. The importance of the first 80 ms of room impulse responses: A historical review.　John Bradley　(Natl. Res. Council, Montreal Rd., Ottawa K1A 0R6, Canada)

The direct sound and reflections arriving within 80 ms after it are particularly important to the perceived quality of sound in rooms such as concert halls and auditoria and have been the focus of the majority of research on concert hall acoustics over the past 50 years. Most of the sound energy is usually concentrated in this time interval and this early sound energy determines many of the perceived characteristics of the hall. Although some remarkable early experiments date back to the 1850s, the major advances concerning how these early reflections are perceived have occurred more recently. Monaural effects including explanations of how our hearing system integrates early reflections, the precedence effect, and the development of measures of clarity and intelligibility have developed from initial work in the 1950s. A little more recently the binaural effects leading to perceived spatial impression have been explained and shown to be critical to achieving good concert halls. Physical measures have been developed that relate to both the monaural and binaural subjective phenomena. This work will review the development of our understanding of the importance of early arriving sound on the acoustical characteristics of concert halls and auditoria.

1:30

1pAA2. Early reflection energy in concert halls: how much, how early, and from where.　Anders Chr. Gade　(Section of Acoust. Tech., Oersted DTU, Tech. Univ. of Denmark, Bldg. 352, DK 2800 Lyngby, Denmark, gade@dat.dtu.dk)

Today, the importance of distributing early reflection energy to listeners and performers in concert halls is well understood and accepted—also among architects. Still, implementation in the practical design of a large hall is not easy, partly because we still have difficulties quantifying precisely the demands in detail (how much, how early, and from where) due to limitations in our acoustic knowledge—and often also due to little response from the clients regarding their preferences when decisions between alternative demands must be made. Therefore, in general it is up to the acoustician to decide how far to promote the good cause on the basis of his/her experience, taste, and talent in influencing the decision process. The aural presentation will focus on the current limitations in our knowledge regarding the musicians' need for early reflections, which is a special challenge in the design of terraced, ''surround'' concert halls. In the case of the listeners, knowledge has reached a much higher level, and recent results regarding how room shape influence the objective acoustic parameters monitoring the early energy will be presented.

1:55

1pAA3. ''REC Curve'' data measured in halls opened after 1988.　Yasuhisa Toyota, Keiji Oguchi, and Minoru Nagata　(Nagata Acoustics, Inc., 201 Ocean Ave., Ste. 1205B, Santa Monica, CA 90402, toyota@nagata.co.jp)

As the results of the studies on early reflections in auditoriums, ''Reflection Energy Cumulative Curve excluding direct sound'' (REC Curve) was introduced and defined at the ASA Hawaii meeting in 1988. In the study, early reflections up to around 80 ms after direct sound were found to be substantially important and influenced by the room shape of the auditorium. Some concert halls were designed based on ''REC Curve'' so that REC Curve up to 80 ms showed smoothly increasing curves in the auditoriums. ''REC Curve'' data measured in the halls opened after 1988 are reported in this paper with some subjective comments on the halls.

2:20

1pAA4. The nervous spatial behavior of early-energy measures in auditoria.　Diemer de Vries and Edo M. Hulsebos　(Lab. of Acoust. Imaging and Sound Control, Delft Univ. of Technol., P.O. Box 5046, 2600 GA Delft, The Netherlands, diemer@akst.tn.tudelft.nl)

Most commonly used room acoustical measures are defined as the ratio between the energies in two time windows of the impulse response. In the Clarity Index and the Lateral Energy Fraction—supposed to be indicators for transparancy and spaciousness, respectively—the time window 0–80 ms plays a predominant role. When impulse responses are recorded along an array of closely spaced microphone positions and the above measures are calculated for each position, it appears that large fluctuations, up to a factor

of 2, occur on a small spatial scale, e.g., in front of one and the same seat. These fluctuations are due to interference of early relections and can be physically well understood. It makes clear that the acoustic quality of a seat or row should not be based on one single measurement but on some spatial average, or that the measures should be redefined such that interference, to which our hearing mechanism apparently is not sensitive, is eliminated.

2:45

1pAA5. Edge diffraction and scattering in the early room impulse response. Rendell R. Torres, Mendel Kleiner (Chalmers Rm. Acoust. Group, Chalmers Univ. of Technol., SE-412 96 Gothenburg, Sweden), and U. Peter Svensson (Dept. of Telecomm., Acoust. Group, Norwegian Inst. of Technol., NO-7491 Trondheim, Norway)

Although reverberation is a common and important measure of quality in architectural acoustics, the early part of a room's impulse response is equally critical, as its spectrum and transient structure predominantly determine the initial filtering (i.e., the natural coloration) that a room endows on an (otherwise anechoic) source signal. Since this initial room filtering could make the difference between an acoustical impression that is, for example, either ''clear and defined'' or ''murky and unfocused,'' various studies have been performed to better understand and model the effects of edge diffraction and surface scattering on the early reflections in room impulse responses. These studies include auralizations with varying ''diffusion'' coefficients, computations and measurements of edge diffraction in early room impulse reponses, and application of ''boss models'' to compute hemispherical scattering in room acoustics auralization. Various results show that these effects can be clearly audible and that their audibility and perceived character depend on the input signal, among other factors. Conclusions are complemented with suggestions for future work. [Work supported by the Axson Johnson Foundation and Teknikbrostiftelsen, Sweden.]

3:10–3:25   Break

3:25

1pAA6. Effects of seating arrays on subjective quality in auditoria. Masayuki Morimoto, Sachio Hirose (Environ. Acoust. Lab., Kobe Univ., Nada, Kobe 657-8501, Japan), and Koji Ishida (Onosokki Co., Ltd., Midori, Yokohama 226-8507, Japan)

Listening tests were performed to clarify the subjective effects of seating arrays. First, a qualitative assessment was made of the subjective effects introduced by the seating array. Test signals were prepared by convolving dry music with impulse responses measured in a scale model consisting of a stage, a floor with audience but no walls or ceiling. The measured impulse responses showed that the seating array caused an increase in amplitude below 80 Hz as well as a large dip at 150–200 Hz. The results of listening tests demonstrated that the seating array affected not only bass perception but also loudness. Second, it was investigated whether reflections from the ceiling and walls can compensate for an excess or deficiency of perceived bass caused by the seating array. In addition, the relation between the frequency characteristics of sound fields and bass perception is discussed. Test signals were prepared by adding a ceiling reflection and reverberation to test signals used in the first test. The results indicated that to a certain extent they can compensate for an excess or deficiency of perceived bass. It is suggested that bass perception should be evaluated according to the frequency characteristics of sound fields within the first 100 ms or so after the direct sound.

3:50

1pAA7. Sound coloration from (very) early reflections. Tor Halmrast (Statsbygg, P.O. Box 8106, Dept. N-0032, Oslo, Norway)

Coloration is defined as changes in Timbre/Klangfarbe. Adding a reflection automatically changes the frequency response of a signal, giving some kind of coloration. This might be looked upon as distortion. However, reflections have been a part of sound distribution since the Greek amphitheatres, indicating that some coloration must be acceptable, even wanted. The question is: Which reflections give disturbing/unwanted coloration? Most room-acoustic criteria assume that everything happening before/after a certain time is supposed to be good/bad (e.g., 50 ms/80 ms for speech/music). We need to take a closer look at the distribution of reflections within these time intervals, to investigate coloration. This paper will give results from measurements in concert halls (Oslo, Munich, Vienna), and opera houses, and compare them with psychoacoustic studies on coloration. It is shown that measurements should be done with the orchestra present on the stage/in the orchestra pit (TOR, through orchestra impulse response). Lonesome/strong reflections with a time delay about 10 ms will give box-klang-coloration. On the other side, colorating reflections might be useful and fun, used in the right way, to support bass instruments, etc. [J. Sound Vib. **232** (2000)].

Contributed Papers

4:15

1pAA8. An investigation of the spectral effect of multiple early reflections. Alexander U. Case (Sonics in Architecture, Rensselaer Polytechnic Inst. and Cavanaugh Tocci Assoc., Consultants in Acoustics, 327F Boston Post Rd., Sudbury, MA 01776, acase@cavtocci.com)

The interaction of a direct sound with a single reflection of short delay is well-understood. Comb filtering is a quantifiable, measurable, inevitable result. In actual auditoria, one rarely hears a single early reflection. Typically a volley of several reflections follows the direct sound. This paper presents an analytical evaluation of the change in frequency response due to, not one, but several reflections. The combination of a direct sound with several early reflections does not result in a more pronounced, more audible comb filter quality. Rather, the interaction of the direct sound with

each of the several reflections in combination with the interaction among the many reflections themselves tends to average out or randomize the expected, comb filtered, spectral distortion.

4:30

1pAA9. When all sound is mere reflections. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301)

For many decades in the recent past, impulse response measurements have been performed in music and drama spaces to assess the clarity of sound and ease of communication offered on a stage or within the audience chamber. Much of the focus has been on the strength and timing of early sound reflections arriving within 80 or so ms following the direct sound. Generally, the thought and design process revolve around an un-

occluded, straight-line path between source and receiver. This paper, however, presents the results of numerous studies and completed buildings in which excellent communication has been achieved in drama and music conditions where the direct sound is attenuated due to source directivity factors and/or obstructions normally encountered within rooms. Results of computer model studies, photos of completed spaces and acoustic data will be presented along with general conclusions.

4:45

1pAA10. The acoustics of the church San Francisco of Santiago of Chile. Jaime Delannoy (Dept. of Acoust., Tech. Univ. Vicente Perez Rosales, Brown North 290 Nunoa, Santiago, Chile, jaime.delannoy@uvipro.cl)

The church San Francisco, declared a national monument, is the most important representative of the colonial architecture that remains in Chile. The definitive construction data of 1618, and after supporting to earthquakes and fires, conserves its original central shape to the present time. The interest of diverse musicians to interpret its works in this enclosure was the reason for the accomplishment of this study. Measurements of the following objective parameters were made: EDT, C80, C50, D50, U50, ti, and NC. The results indicate that a high background noise exists (NC=46) and a poor intelligibility (D50 lower 0.34). Nevertheless its inner distribution provides an initial favorably short time delay gap (ti=22 ms) that provides an acoustic privacy that favors the execution of small orchestras. [Work supported by Department of Acoustics–Technical University Vicente Perez Rosales.]

5:00

1pAA11. The acoustic design of the Concertgebouw, Amsterdam, and resolution of the hall's early acoustical problems. Pamela Clements (Kirkegaard Assoc., 801 W. Adams St., Chicago, IL 60607)

From the outset, acoustical considerations were a driving force in the design of the Concertgebouw. Several halls were used as acoustical models, including the Tonhalle in Düsseldorf, the Leipzig Gewandhaus, and the concert room of the Felx Meritus Society in Amsterdam. When the Concertgebouw opened in 1888, however, it had acoustical problems: the newly formed orchestra struggled with the hall's long reverberation time, and balance between brass and strings was difficult. Various remedies were tried, including potted plants to ''soak up'' the sound. The installation of the organ in 1891 was thought by some to make the acoustics worse. Between 1895 and 1900 the problems were gradually resolved. Orchestral balance was improved in 1899 when the orchestra platform was rebuilt to reduce the steepness of the risers. Other factors included the approach of the new conductor, Willem Mengelberg, improved performance standards, and the arrival of the late Romantics, particularly Tchaikovsky, in the repertoire. Happily this coincided with improved transportation that enabled newly enthusiastic audiences to fill the hall. The paper uses plans, photographs and other historical evidence to explore the intent of the acoustic design of the Concertgebouw, and its transformation from near failure to musical and acoustical success.

MONDAY AFTERNOON, 4 JUNE 2001 ADAMS ROOM, 1:30 TO 4:55 P.M.

Session 1pAO

Acoustical Oceanography: Acoustical Instrumentation for Water Column Measurements II

Kenneth G. Foote, Chair

Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

Chair's Introduction—1:30

Invited Papers

1:35

1pAO1. Wideband and correlation techniques and their application to fisheries acoustics: Existing prototypes and future trends. Manell E. Zakharia (Ecole Navale, French Naval Acad., Batiment Bougainville, 29240, Brest Naval, France, zakharia@ecole-navale.fr)

The aim of this work is to describe several ''research sonar'' prototypes (one-of-a-kind), to show the performance improvement by using both wideband techniques and associated echo processing procedures such as correlation, synthetic aperture, or spectral signature analysis. We will describe the improvement of fisheries sounders provided by widening the frequency range: enhanced axial resolution, smoothed spatial and angular variations. The advantages will first be illustrated by experiments on individual fish. Then, the performance of such systems will be investigated at sea using wideband prototypes. A wide range of processing algorithms will then be detailed: pulse compression, dynamic focusing, beam control on a wide frequency range, synthetic aperture, echo to echo information, etc. Some of them have been commonly used in military sonar systems, some others are still under exploration in other areas (echo–echo correlation, synthetic aperture processing of sidescan data). A multibeam approach based on wideband correlation will also be described. Finally, the performance of wideband classification of fish species at sea will be shown. The technical limitations (transducers, acoustic compatibility, real-time processing, etc.) that could limit the development of such systems will be developed and several solutions will be proposed to overcome the development complexity.

1pAO2. A synopsis of state-of-the-art high-frequency sonars for water column measurements. Mark V. Trevorrow (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada, Mark.Trevorrow@drea.dnd.ca)

In recent years there has been a dramatic surge in the use of high-frequency (10 kHz–2 MHz) backscatter sonar systems for oceanographic and bioacoustic measurements in oceans, lakes, and rivers. These applications have generally focused on measuring the distribution of scatterers, such as bubbles, particulates, microstructure, fish, and zooplankton. Additionally, oceanographic flows and wave motions can be illuminated using these scatterers as tracers, using both the intensity and Doppler signals. Advances in electronics, data storage, signal processing, and transducer materials technology have made, and will continue to make, HF sonars smaller yet more capable in terms of bandwidth, resolution, and endurance. Some novel applications of simple echo sounders and sidescan sonars will be reviewed, showing examples of near-surface bubble distributions, internal hydraulic flows, and fish and zooplankton detection. Newer developments in multifrequency, volumetric multibeam and parametric sonars will be discussed.

3:05–3:25 **Break**

Contributed Papers

3:25

1pAO3. Expanding capabilities of acoustic instruments: Benefits of calibrated output signals, bandwidth, midwater signals, and dynamic range. Kenneth G. Foote (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Echo sounders, multibeam sonars, acoustic Doppler current profilers (ADCPs), and other acoustic instruments are widely used, if often for quite specific applications. Sometimes, these applications have been pursued at the design stage to the detriment of general system performance and versatility. Availability of a calibrated output signal in a fishery echo sounder, for example, may enable this to be used in quantitative applications. The availability of bandwidth may enable scientific echo sounders to classify and quantify multispecies assemblages of zooplankton. The provision of midwater signals may enable multibeam bathymetric sonars to be used for imaging scatterers in the water column. Extension of the dynamic range of ADCPs may enable a single unit to be used as four synchronous echo sounders with noncollinear beams. In general, the capacities of most underwater acoustic systems can be expanded significantly by attention to issues such as those mentioned here. Resulting opportunities for new applications are described, with justification in the physics of scattering. [Work supported by the Alfred P. Sloan Foundation.]

3:40

1pAO4. Multibeam SONAR from the bottom and up. Jeff Condiotty (SIMRAD, Inc., 19210 33rd Ave. W., Lynnwood, WA 98036) and Oywind (Irv) Bjorkheim (Kongsberg Simrad, Inc., Lynnwood, WA 98036)

Acoustic multibeam technology has traditionally been used for hydrographic surveying but is now being introduced as a useful tool in data collection within the water column. A new PC-based system operating at frequencies of 90 or 200 kHz allows visualizing targets within the water column for fisheries behavioral and enumeration studies as well as bottom characterization and profiling, at the same time. The system takes advantage of a wide sampling swath, up to 150 with up to 128 individual beams. The system can be operated in a wide beam mode or used in a Mill's Cross mode to control the beamwidth in two dimensions. Data are displayed as a SONAR image in real time while raw data (pre-beamformed or beamformed) can be stored for further processing utilizing commercial software. The transducer can be either hull mounted, installed on a submersed vehicle, or attached to a pole mount on a small boat. Applications and data collection methodology to visual and measure target reflectors will be discussed along with calibration of individual beams to produce quantitative results.

3:55

1pAO5. Target strength analysis in high fish densities using a probing split beam transducer. Egil Ona and Ingvald Svellingen (Inst. of Marine Res., P.O. Box 1870, 5024 Bergen, Norway)

Unbiased measurements of the mean target strength are needed if the absolute fish abundance is to be measured. For most pelagic fish, this also requires direct target strength techniques to work in high-density situations. A probing, pressure stabilized split beam transducer has been lowered from the research vessel directly into dense herring layers in order to resolve a single target at short range. Specialized hardware has been used for single fish target tracking, from which data on target strength, target directivity, swimming speed and tilt angle are derived. The paper will focus particularly on calibration procedures for transducer depth performance, target-tracking methodology and on the monitoring of essential transducer movement parameters.

4:10

1pAO6. Instantaneous target strength measurement of live squid with synchronized video image. Tohru Mukai and Kohji Iida (Grad. School of Fisheries Sci., Hokkaido Univ., 3-1-1 Minato-cho, Hakodate, Hokkaido 041-8611, Japan, mukai@fish.hokudai.ac.jp)

The TS is a complex function of the fish's size, shape, orientation, behavior, and internal structure, as well as of the acoustic wavelength employed. Therefore, for practical purposes, it is necessary to measure the TS experimentally. There is a considerable amount of knowledge about the TS of fish. Information about the TS of squid, however, is much less. Through previous studies over the past 6 years at our experimental field, acoustic equipment has been modified and gradually stabilized. It was also possible to use live squid in our specially designed experimental barge equipped with two underwater video cameras which look at the side view and upward. Preliminary TS and orientation measurements were done on single tethered live squid *Todarodes pacificus*. The live squid was suspended with monofilament nylon string near the acoustic beam axis and positioned 3.5 m below the transducers. The echo signals of the squid at six frequencies were converted to amplitude modulated signals in audible frequency, then they were recorded on audio channels of video tape with a video image simultaneously. Afterward, the echo envelopes were demodulated to analyze the relationship between TS and the swimming aspect of live squid by synchronized video image. Preliminary results from this measurement will be presented.

4:25–4:55

Panel Discussion

Session 1pEA

Engineering Acoustics: Hydrodynamics in Ducts, Jets, and Blowers

Daniel M. Warren, Chair

Knowles Electronics, 1151 Maplewood Drive, Itasca, Illinois 60143

Chair's Introduction—2:00

Contributed Papers

2:05

1pEA1. Determination of fan inlet duct mode structure using in-duct microphone array measurement. Carl H. Gerhold (MS 461, NASA Langley Res. Ctr., Hampton, VA 23668, c.h.gerhold@larc.nasa.gov)

A methodology is developed in which the amplitudes of propagating modes are extracted from the sound measured in the duct. The method is derived from a beamforming technique in which the weighting function is the wave description of a chosen mode. The modal distribution that arises from the analysis is used to synthesize the sound propagating in the duct which is projected to the far field to estimate the sound radiated from the duct inlet. The methodology is tested using an array of dynamic pressure transducers in the inlet duct wall of a turbofan engine. The data show that the modal structure of the blade passage frequency tone at lower fan speeds is predominately from the interaction of the rotor with downstream obstructions. At higher speeds, when the fan blade tip speed becomes supersonic, many modes are excited. Good agreement is found between measured and estimated far field sound levels in the forward quadrant of the engine at the lower engine speeds. Significant divergence between measured and calculated far field sound occurs when the blade tip speed becomes supersonic. The methodology is a useful identification of the noise generation mechanisms present in the fan.

2:20

1pEA2. Proof of existence of monopole-type sound generation mechanism in a centrifugal blower. Ashnapreet Nagi and Sean Wu (Dept. of Mech. Eng., Wayne State Univ., 5050 Anthony Wayne Dr., Detroit, MI 48202)

Previous studies have shown that sound radiation from a centrifugal blower is primarily of the dipole and quadrupole characteristics, which describe the effects of fluid-structure interaction and turbulence generated by rotating blades. However, in many engineering applications the sound powers from centrifugal blowers are found to be proportional to the speed to the power of 4–6, which implies the existence of a monopole-type sound according to the Lighthills acoustic analogy [Lighthill, Proc. R. Soc. London, Ser. A **222**, 564–587 (1952)]. This paper demonstrates that such a monopole sound generation mechanism indeed exists in a centrifugal blower used in an HVAC unit of a passenger vehicle. LDV measurements show a velocity fluctuation at the trailing edge tangential to the blade surface of an impeller. Laser tachometer measurements further indicate a fluctuation in shaft speed for all models of blowers tested. The resulting sound powers from these blowers show a predominant dependence of the speed to the power of 4, indicating the characteristic of a monopole. When these blowers are installed on a nonfluctuating shaft, however, the flow fluctuations are significantly reduced and the resulting sound powers are found to be proportional to the speed to the power of 6, which implies that the monopole sound generation mechanism is effectively eliminated.

2:35

1pEA3. Radial temperature variations in cylindrical waveguides and implications for flow measurement. Morten Willatzen (Mads Clausen Inst., Univ. of Southern Denmark, Grundtvigs Alle 150, DK-6400 Sonderborg, Denmark)

A quantitative treatment of radial temperature variations in a cylindrical duct on ultrasonic flow meter performance is discussed in the laminar flow regime. First, based on the continuity equation, the Navier–Stokes equations, and an energy equation including loss mechanisms due to heat conduction and viscous effects, the steady-state temperature and flow spatial distributions are determined in two cases of practical interest: (a) cylinder wall temperature is maintained at a constant value, and (b) cylinder wall temperature decreases linearly with cylinder axial distance. It is shown that while radial temperature variations are insignificant in case (a), radial temperature gradients as large as 100–200 K/m are possible in case (b) for a fixed axial temperature gradient decrease of 0.1 K/m. Such strong temperature gradients have strong and unfortunate consequences for flow measurement. Large flow meter errors—up to several percentages—are possible using typical parameter values for water as medium. Finally, it is shown that effective ways exist such as to diminish the influence of temperature gradients on flow meter performance. Besides the obvious choice of insulating the flow meter tube, flow measurement errors due to radial temperature variations can be effectively suppressed by reducing the cylinder radius and/or ultrasound frequency.

2:50

1pEA4. Computational investigation of acoustic streaming in a two-dimensional duct. W. R. Erskine and A. Selamet (Dept. of Mech. Eng., Ohio State Univ., Columbus, OH 43210)

The present work investigates the nonlinear phenomenon of Rayleigh streaming characterized by acoustically driven flows attenuated by wall friction, resulting in nonzero average streaming velocities and circulation zones within the duct. A two-dimensional, long and narrow duct is considered in the present study with one end closed and the other subject to a sinusoidal acoustic velocity of varying frequency and amplitude, leading to a standing wave within the duct. A computational fluid dynamics code is used to solve the unsteady, compressible Navier–Stokes equations in two dimensions with turbulence modeling. Since the formation of streaming depends on the velocity profile within the Stokes layer, a high resolution grid is employed near the walls. The effect of inlet conditions is examined in terms of the circulation patterns, including the number of zones per wavelength and penetration depth into the fluid, as well as the mean velocity profiles.

3:05–3:20   Break

3:20

1pEA5. On the sound sources of the screech tone radiated from a choked circular jet. Yoshikuni Umeda and Ryuji Ishii (Yoshida Honmachi, Sakyo-ku, Kyoto, Japan, umeda@kuaero.kyoto-u.ac.jp)

In this experimental study, the generation of the screech tone in the dominant helical oscillation mode C was investigated by using a series of instantaneous Schlieren photographs and it was observed that the moving sound source for mode C appears to be the prominent point on the helical-shaped coherent vortical structure. This sound source is generated by the interaction between the helical-shaped vortical structure and the oblique shock at the rear edge of the third shock cell. The sound source rotates about the jet axis in the plane perpendicular to the jet axis at a constant supersonic speed and forms a Mach cone. The envelope of the Mach cone widens upstream and downstream as the sound source moves along the circular orbit and forms two helical-shaped wave-fronts of the screech tone propagating in upstream and downstream directions.

3:35

1pEA6. The multiple source feedback of supersonic jet screech. IV. Space–time diagrams and the fast/slow waves of transitory wave theory. Alan Powell (Dept. of Mech. Eng., Univ. of Houston, Houston, TX 77204-4792)

In Part II [J. Acoust. Soc. Am. **108**, 2588 (2000)] the space–time wave diagrams of multiple (point) source screech were elucidated, showing the condition for perfect, or nearly so, reinforcement of the feedback sound at the nozzle. In the transitory wave theory [J. Acoust. Soc. Am. **106**, 2194 (1999)] the velocity modulation of a transitory instability wave as it traverses the periodic cell structure results in a *continuous* phased array which was resolved into fast (approx. sonic) upstream and slow downstream source waves (of constant sinuous amplitude, apart from the envelope factor), only the former radiating strongly, in the upstream direction. The two points of view are reconciled by constructing fast and slow waves in the space–time diagrams by: (a) introducing intermediate negative sources: (b) taking the sources at half strength and adding virtual sources at the zero positions on the fast wave traces, of the same sign as adjacent sources: (c) taking the other half-strength sources and adding virtual sources of opposite sign, the sources then being of the same sign along the slow wave traces. The implications of the need to introduce negative sources in the point source array are discussed.

3:50

1pEA7. Vortex shedding characteristics of two cylinders with differing diameters. Muhittin Sylemez (Dept. of Ocean Eng., Istanbul Tech. Univ., Ayazaga Kampusu, 80626 Istanbul, Turkey) and Mehmet Atlar (Univ. of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK)

This paper presents an experimental investigation of the vortex shedding characteristics of two circular cylinders with differing diameters in representing a simple riser configuration. The smaller of the two cylinders is located at varying positions around the circumference of the larger cylinder. Measurement of forces on the cylinders and the associated shedding frequencies are recorded and analyzed for five different flow velocities and orientations of the riser in the Emerson Cavitation Tunnel of Newcastle upon Tyne University. The Reynolds number corresponding to the test conditions varies from 9.64×10^4 to 4.82×10^5 for the large diameter of the cylinder and from 3.5×10^4 to 1.75×10^5 for the small diameter cylinder. The experiments were split into two distinct phases: the first was flow visualization using the hydrogen bubble technique and the second phase was force and shedding frequency measurements.

MONDAY AFTERNOON, 4 JUNE 2001

PDR 18, 1:00 TO 2:15 P.M.

Session 1pPA

Physical Acoustics: Nondestructive Testing

David H. Chambers, Chair

Electrical Engineering, Lawrence Livermore National Laboratory, L-154, P.O. Box 808, Livermore, California 94551-5508

Contributed Papers

1:00

1pPA1. Propagation of Lamb waves in a bonded plate. Xiao-Min Wang and Ming-Xuan Li (Inst. of Acoust., Chinese Acad. of Sci., 17 Zhongguancun St., Beijing 100080, PROC, xiaomwang@2911.net)

A solution to the propagation of Lamb waves in a two-layered elastic plate for a variety of interfacial conditions is obtained. The bond rigidity between the adjacent layers is accounted for by a spring model proposed by many researchers. Instead of being in the conventional form of a determinant, the resulting dispersion equation is given in a form of polynomials with the normal and tangential interfacial stiffness constants as the sum-coefficients. This novel expression of dispersion equation has an explicit physical interpretation and application for inversion of the stiffness constants. The influences of the variations of the stiffness constants on the propagation of Lamb waves are numerically studied. The results for different adhesive states show that the degradation of the bond rigidity may cause the decreasing velocity and the frequency-shifts of a given Lamb wave mode, thus having a possibility for inspecting adhesively bonded joints by using the ultrasonic wave velocity measurements. Experimental results to verify the theoretical analysis are given and the empircal relationship between the bond rigidity and the Lamb wave velocity of several lower modes is presented. [Work supported by the National Natural Science Foundation of China.]

1pPA2. Ultrasound propagation in random, two-phase media with application to damage detection in concrete. Joseph A. Turner, Sean P. M. Dugan, and Liyong Yang (Dept. of Eng. Mech., W317.4 NH, Univ. of Nebraska–Lincoln, Lincoln, NE 68588, jaturner@unl.edu)

The scattering of ultrasonic waves in a random two-phase medium, such as concrete, is discussed. Explicit expressions are derived for the attenuation of longitudinal and shear waves in terms of the statistics of the density and Lamé parameter fluctuations. The derivation is based upon diagrammatic methods. The problem is posed in terms of the Dyson equation, which is solved for the mean field response. The attenuation results are given here in a straightforward manner. The general description is further simplified in accord with assumptions representative of a random, two-phase material. The medium is assumed statistically homogeneous and statistically isotropic, with a single correlation length. With these assumptions, the attenuations reduce to simple integrals on the unit circle. Expressions for mean free paths and elastic diffusivity follow from the general scattering theory. Applications to diffuse ultrasound experiments in concrete are then discussed. In particular, the appropriate form of the two-point spatial correlation function is examined and compared with experimental results from digitized concrete images. These results show that the assumed form of the correlation function describes the spatial statistics well. Finally, damage detection prospects relevant for concrete inspection are discussed. [Research supported by NSF Grant No. CMS-99-0078707.]

1pPA3. A crack model for nonlinear wave modulation spectroscopy. Boris Gommerstadt (Northeastern Univ., Boston, MA, 02115 Gommers@coe.neu.edu) and Alexander Sutin (Stevens Inst. of Technol., Hoboken, NJ 07030)

Nonlinear wave modulation spectroscopy is a novel technique for damage and crack evaluation based on the effect of modulation of ultrasonic waves by vibration. This technique was developed as an experimental approach and our goal is to discuss some mathematical models simulating these phenomena. The first model describes variation of resonance frequencies of the ultrasonic wave by breathing crack oscillation under low frequency vibration. This model considers an opened crack with the length changing under the action of vibration. The second model is concerned with plastic deformation in the vicinity of the crack tips that develops a hysteretic stress–strain relationship leading to the increase of ultrasonic wave damping. The Dugdale model is employed to describe ultrasonic wave damping and its variation under the vibration. These models relate modulation parameters to crack size and position. The suggested models explain the basic features of the observed modulation phenomena and could be used for characterization of the dimension and localization of cracks and crack-like damage.

1pPA4. Crack location using nonlinear means applying modulation of ultrasonic pulses by cw vibration. Vyacheslav V. Kazakov (Inst. of Appl. Phys. RAN, Nizny Novgorod, Russia), Alexander M. Sutin (Stevens Inst. of Technol., Hoboken, NJ), and Paul A. Johnson (Los Alamos Natl. Lab., Los Alamos, NM)

Nonlinear wave modulation spectroscopy (NWMS) was recently introduced as a new tool in NDE. This technique employs the nonlinear interaction of ultrasound and vibration in the presence of defects. Vibration changes the contact area within a defect effectively modulating an ultrasonic wave sensing the defect. NWMS has high sensitivity to the presence of cracks and can be a very good tool for a ''pass–fail'' test but cannot provide the crack location. A new method for locating defects or cracks in a material, presented in this paper, is based on the modulation of ultrasonic pulses by vibration. The cw vibration induces modulation of the signal reflected from cracks. Measurements of the spatial distribution of the modulation level are the basis of nonlinear acoustic imaging. Experimental verification of the method has been conducted in steel plates containing a circular hole and a fatigue crack. The high-frequency (carrier) impulse was about 3 MHz, and the modulating vibration frequency was 5–15 Hz. This technique images a crack very well while a hole was not detected. As the location of damage is the next important issue in NWMS, this work represents an important step. [This work was partially supported by International Science Technical Center, Grant No. 1369.]

1pPA5. Laser-induced weak shock in fused silica for destructive testing of thin films. Junlan Wang and Richard L. Weaver (Theoretical and Appl. Mech., Univ. of Illinois, Urbana, IL 61801)

Dynamic testing of the adhesion of thin films to solid substrates is a challenging problem. Work of the past decade testing such films by the impact of a compressive stress wave pulse generated by a laser pulse has determined that a critical parameter is the fall time of the stress pulse. Here we report the development of a variation that allows the necessary fall times to be readily achieved. A YAG laser of moderate power (100 mJ) and pulse duration (10 ns) illuminates a 0.4-μm aluminum layer on a fused silica substrate. A transparent 300-μm water–glass layer provides an inertial backing. The resulting compressive stress pulse that enters the substrate has a strain amplitude of the order of 2% and a duration of 10 ns. The negative elastic nonlinearity of the fused silica causes this pulse to evolve, over a distance of 1 mm, into a linear ramp followed by a weak shock that has a theoretical fall time of 10 ps. Interferometric measurements of the pulse, and observations of the failure of a micron-thick aluminum layer on the opposite side of the substrate, confirm this understanding. [Work supported by NSF.]

Session 1pPP

Psychological and Physiological Acoustics: Exploring the Dynamics of Auditory Perception Through Neuroimaging

Ann Clock Eddins, Chair

Center for Hearing and Deafness, University of Buffalo, 215 Parker Hall, Buffalo, New York 14214

Invited Papers

1:30

1pPP1. ''What'' and ''where'' in the auditory cortex. Josef P. Rauschecker (Georgetown Univ. Medical Ctr., WP15 NRB, 3970 Reservoir Rd., N.W., Washington, DC 20007)

Several cortical fields have been defined in the superior temporal gyrus (STG) of the rhesus monkey. These ''belt'' areas are further characterized by their responses to spatially localized, species-specific vocalizations. Neurons in the caudal belt are more specific for spatial location, whereas neurons in the rostral belt are more selective for communication calls. Furthermore, the caudal and rostral belt project to distinct regions of the prefrontal cortex previously identified as subserving roles in visual spatial analysis and object recognition, respectively. Functional neuroimaging in humans also supports the existence of dual streams for the processing of sound content and location. Complex sounds generally activate a larger area in the belt region of the STG than tones. Consonant–vowel combinations give rise to activity distributions anterior to Heschls gyrus, which are organized into ''phonetic maps.'' If visual input is added in a McGurk-type fusion experiment, these maps are found to shift in individual subjects. Virtual auditory space stimuli administered via headphones activate two regions in the posterior parietal cortex. When this activation is compared with that by a visual spatial task, no overlap is found in the inferior parietal cortex, whereas convergent input is present in a superior parietal region.

2:00

1pPP2. Functional imaging of temporal auditory processing. Timothy D. Griffiths (Functional Imaging Lab., Queen Square, London WC1N 3BG, UK, t.d.griffiths@fil.ion.ucl.ac.uk)

Functional imaging of the auditory system can be carried out using positron emission tomography (PET) and functional MRI (fMRI). Both techniques depend on the blood flow response to stimuli as a measure of brain activity, a process with a delay of 10 s in the auditory cortex. This limits the ability of the techniques to investigate temporal processing in the auditory system. Experiments to investigate the basis for temporal processing are critically dependent on the model that is used to interpret the brain blood flow data. The blood flow response in such experiments depends on local mean synaptic activity; this can be increased either by the firing rate in a subpopulation of cells or by increased synchronization of neural activity. At the millisecond level, an increase in the fMRI BOLD response with the temporal regularity of a signal at the millisecond level has been demonstrated as early as the cochlear nucleus. This may be due to an increased firing rate of onset chopper cells tuned to a particular autocorrelation delay or to increased synchronized activity in the whole population. The temporal processing of patterns over longer periods (seconds) is likely to involve areas of the cortex distinct from the primary auditory cortex.

2:30

1pPP3. Cognitive and sensory influence on the perception of complex auditory signals. Ann Eddins, David Eddins (Dept. of Communicative Disord. and Sci., Univ. at Buffalo, Buffalo, NY 14214), Mary Lou Coad, Alan Lockwood (School of Medicine, Univ. at Buffalo and WNY VA Medical Ctr.), and Charles Watson (Indiana Univ.)

Using positron emission tomography (PET), we have been studying the relationship between psychophysical performance and cortical activation during auditory discrimination tasks. In one experiment, we studied normal hearing listeners who had markedly different perceptual performance on a battery of auditory discrimination tasks. Changes in regional cerebral blood flow (rCBF) were measured while subjects discriminated changes in a tonal sequence or a temporally analogous syllable sequence. The results showed that behavioral performance was correlated with cortical activation in auditory sensory areas as well as cognitive areas related to attention and memory, depending on the demands of the task. To evaluate further whether attention to specific features of complex auditory signals influences the recruitment of cognitive processing areas of the cortex, we studied changes in rCBF when subjects

2309   J. Acoust. Soc. Am., Vol. 109, No. 5, Pt. 2, May 2001

141st Meeting: Acoustical Society of America   **2309**

1p MON. PM

attended to different features of a 1-octave band signal that varied in spectral envelope frequency (i.e., ripple frequency) or carrier frequency. The results showed that when subjects attended to ripple frequency relative to carrier frequency, activation in the thalamus increased. Moreover, when they discriminated ripple frequency in a low-carrier versus a high-carrier frequency, there was increased activation in the anterior cingulate gyrus, which is known to be involved in attention function.

3:00–3:15   Break

3:15

1pPP4. Some neural correlates of speech and nonspeech perception. J. R. Binder, E. Liebenthal, J. N. Kaufman (Dept. of Neurology, The Medical College of Wisconsin, 9200 W. Wisconsin Ave., Milwaukee, WI 53226), R. L. Piorkowski, and R. E. Remez (Barnard College, New York, NY 10027)

In a series of functional magnetic resonance imaging experiments aimed at examining the perceptual boundary between speech and nonspeech stimuli, brain activation signals were measured while subjects listened passively to a continuum of sounds ranging from noise to speech, detected speech signals in varying levels of noise, and perceived the same sinewave speech replicas as either nonspeech or speech while performing an auditory discrimination task. Activation at the anterolateral aspect of left Heschl's gyrus increased with signal-to-noise ratio in the passive listening and noise masking experiments, and was higher when subjects heard sinewave speech as nonspeech than when they heard the same stimuli as speech while performing an auditory task that conflicted with the speech percept. Activation in this area was correlated with behavioral measures of phoneme perception and was unrelated to attentional demands. The data suggest a key role for this region in the representation of spectrotemporal information underlying phoneme recognition.

3:45

1pPP5. The functional anatomy of tinnitus. Alan H. Lockwood (VAWNYHS and Univ. at Buffalo, Buffalo, NY 14215), Richard J. Salvi, Robert F. Burkard, and Samuel A. Reyes (Univ. at Buffalo, Buffalo, NY 14215)

We have used positron emission tomography to map the neural systems that mediate subjective tinnitus. In two studies, subjects altered their tinnitus by voluntary jaw muscle contraction or sustained lateral gaze. In the third, tinnitus was altered by intravenous lidocaine. These experiments all linked the loudness of tinnitus with unilateral spontaneous neural activity in auditory cortex. This unilaterality contrasts with the bilaterality of cortical responses during acoustic stimulation, evidence for a noncochlear central origin for tinnitus. Plasticity in the central auditory system, thought to be associated with deafferentation, was demonstrated in hearing-impaired or unilaterally deaf patients by the greater extent of auditory cortex activation by tonal stimuli when compared to normals. Other sensory and motor systems may interact with these aberrant pathways, further upsetting the balance of auditory system activity. This is manifested by patients' reports of transient voluntary loudness and/or pitch control. In controls, lateral gaze suppressed auditory cortical activity: this cross-modal inhibition was absent in the patients. Subjective tinnitus appears to be due to spontaneous neural activity in aberrant central auditory pathways that have abnormal interactions with other sensory and motor systems. This complexity may explain the unresponsiveness of tinnitus to pharmacological interventions.

Contributed Paper

4:15

1pPP6. Functional imaging of brain activity during speech production. John Sidtis (Dept. of Psych., New York Univ., 6 Washington Pl., 8th Fl., New York, NY 10012)

The functional anatomy of the cerebral cortex, subcortical nuclei, and cerebellum during speech production is understood primarily through the study of defects following brain damage. Functional imaging methods may provide another view of this system in normal as well as in impaired speakers. In the present study, 13 right-handed, native-English-speaking subjects performed a 60-s syllable repetition task during imaging of cerebral blood flow using positron emission tomography (PET). Each subject was scanned eight times (alternating speech and resting control conditions) and their repetitions of the syllables \pa, ta, ka\, produced as quickly as possible, were recorded for acoustic analysis. Blood flow values were measured from 11 regions of interest on each side of the brain. Speech samples were analyzed with respect to several characteristics including rate. Multiple linear regression analyses were used to predict each acoustic measure. Speech rate was predicted by a linear combination of increased flow in the left inferior frontal region (Broca's area) and decreased flow in the right basal ganglia. This functional anatomy is consistent with clinical observations, but was only apparent when performance measures were considered in the analyses. Functional imaging is most useful when both psychological and physiological measures are integrated.

Session 1pSC

Speech Communication: Speech Perception Potpourri (Poster Session)

Joan E. Sussman, Chair

Department of CDS, University of Buffalo, 122 Cary Hall, Buffalo, New York 14214

Contributed Papers

All posters will be on display from 1:30 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 1:30 p.m. to 3:15 p.m. and contributors of even-numbered papers will be at their posters from 3:15 p.m. to 5:00 p.m.

1pSC1. A computational model for detecting and perceiving dynamic properties of speech. Partha Niyogi (Dept. of Computer Sci., The Univ. of Chicago, 1110 E. 58 St., Ryerson 157, Chicago, IL 60637)

We investigate many aspects of a computational framework for detecting phonetically salient dynamic properties of the speech signal from continuous speech input. As a prototypical example of this problem, we consider frameworks for the detection of the transition from closure to burst for stop consonants in continuous speech. We formulate the problem as an optimal non-inear filtering problem. This involves the mapping of the dynamic acoustic data into a low dimension Riemannian manifold in an infinite dimensional Hilbert space. This leads to many kinds of metrics that can now be imposed on the dynamic speech signal and we explore these metrics in the context of the perceptual magnet effect. An extensive discussion of the robustness of the strategy, the nature of the errors made, the relationship to acoustic phonetis and distinctive features and to categorical perception will be conducted. Related problems that have a similar structure include nasal detection using transitions from vowels into adjacent nasals, transitional models of vowels and so on. Results of automatic speech recognition experiments using these dynamic features will be presented.

1pSC2. Auditory context effects for nonequivalent sources. Lori L. Holt (Dept. of Psych., Carnegie Mellon Univ., 5000 Forbes Ave., Pittsburgh, PA 15213) and Andrew J. Lotto (Washington State Univ., Pullman, WA 99164)

One of the most compelling examples of the symmetry in speech perception and production is the ability of the perceptual system to maintain phonetic constancy despite the context sensitivity in acoustic realization of the phonetic segment arising from coarticulation. It may be suggested that the perceptual system is able to accomplish this compensation for coarticulation by recovering the dynamic information in the surrounding acoustic context of the segment. We will present previous data and novel demonstrations that suggest that the contextual information need not come from the same source nor even resemble speech. For example, combinations of pure tones, which are not heard as speech, can shift the identification of neighboring consonants. These nonspeech effects can be described as cases of general auditory contrast. Whereas it is typical in dynamic theories of speech to refer to information specific to vocal tracts in accounting for speech perception phenomena, we feel that the present data are not completely inconsistent with a dynamic approach. The presence of contrast effects suggest to us that the perceptual system respects the dynamic nature of the environment by enhancing sudden canges in kinematics, which are presumed to be caused by application of novel forces.

1pSC3. Perception of English consonants from dynamic acoustic properties. Robert E. Remez, Rebecca L. Piorkowski, Stephanie Wissig, and Abigail Batchelder (Dept. of Psych., Barnard College, New York, NY 10027-6598)

Phonetic perception is often said to depend on short-term attributes of vocally produced sound. Counterevidence to this premise is found in the example of sinewave replicas of speech. The dynamic properties of a sinewave replica evoke phonetic impressions despite the absence of typical or natural sounding short-term elements. To gauge the robustness of phonetic perception from dynamic acoustic properties, we compared identification performance by normal listeners across identification tests composed of signals that preserve dynamic properties of speech despite the absence of short-term speechlike properties: sinewave replicas modeled on the natural samples of 18 English consonants in CV syllables, noise-band vocoded signals derived from the natural samples, and chimerical signals exhibiting the coarse-grain spectrum envelope of natural speech and the fine structure of nonspeech signals. Identification from dynamic acoustic properties was well predicted from control performance using sampled natural speech in noise, showing that short-term acoustic properties and their corresponding auditory qualities are inessential in perceiving consonants. This outcome is consistent with a perceptual account that favors attention to the characteristic modulations of an acoustic carrier, and opposes an account appealing to auditory essences of each phonetic element.

1pSC4. The perception of rate induced resyllabification in English. Kenneth de Jong, J. Byung-jin Lim, and Kyoko Nagao (Dept. of Linguist., Indiana Univ., Bloomington, IN 47405)

Stetson (1951) noted that, when repeated, singleton coda consonants (VC) appear to modulate into onset consonants (CV) as the rate of repetition increases. The current study examines whether nave listeners perceive such resyllabifications, and whether such perceptions are affected by the voicing of the resyllabified consonants. Stimuli were extracted from production experiments in which talkers repeatedly produced singleton voiced and voiceless stops in the CV or VC position. Speakers entrained to a metronome which increased tempo from 450 to 150 ms/syllable. Open-set identification revealed that (1) while slow VCs are identified as such, fast VCs are identified as CVs a majority of the time and (2) CVs are rarely identified as VCs; however (3) both CVs and VCs at fast rates are often identified as CVCs, especially when the consonant is voiceless. A forced-choice identification task indicates that fast VCs and CVs, while clearly differentiated at slow rates, are both identified as CVs 80% of the time at fast rates. These results support Stetson's observations, but like previously reported production results, indicate that fast rate tokens are partially ambiguous between CV and VC forms, an ambiguity which can get resolved by splitting the consonant in question. [Work supported by NIDCD and NSF.]

2311   J. Acoust. Soc. Am., Vol. 109, No. 5, Pt. 2, May 2001

141st Meeting: Acoustical Society of America   **2311**

1pSC5. Dynamic information signals voicing for syllable-final stops.
Susan Nittrouer, Sandy Taylor, Lesley Larive, Carol Manning, and Tammy Mertes (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131)

Several acoustic properties can influence listeners' judgments of voicing for syllable-final stops, including vocalic length. Developmental and cross-linguistic experiments have suggested that listeners with little experience about how vocalic length is related to final-stop voicing weight this property less than more-experienced listeners. For five (English, CVC) minimal pairs, we examined the acoustic consequences of final-stop voicing, as well as listeners' weighting of two properties that signal voicing: vocalic length and syllable-offset transitions. First, an acoustic analysis of words showed that differences in vocalic length associated with final-stop voicing are attenuated when words occur in continuous speech, raising questions about how much experience any listener really has hearing vocalic-length differences correlated with final-stop voicing. Next, adults and children (3, 5, and 7 years old) labeled natural stimuli made from words that originally had voiced or voiceless stops, in which we modified vocalic length. Partial correlations revealed that listeners of all ages weighted offset transitions more than vocalic length. In fact, listeners were reluctant to label any stimulus that did not have transitions clearly indicating closure as ''voiced,'' and so the traditional ''vocalic-length effect'' was apparent only for originally voiced stimuli. We conclude that dynamic information takes precedence in this perceptual decision.

1pSC6. Change deafness: The inability to detect changes in a talker's voice.
Michael Vitevitch (Dept. of Psych., Indiana Univ., 1101 E. 10th St., Bloomington, IN 47405-7007, mvitevit@indiana.edu)

Change blindness is a failure to detect a change in a visual scene. A shadowing task was used to demonstrate an auditory analog to change blindness–change deafness. Participants repeated words varying in lexical difficulty. After a rest-break they heard more words from either the same or a different talker. Answers to explicit questions about the change in talker and implicit measures of behavior (i.e., response latencies) demonstrate that processing is affected by the change, even if participants do not explicitly report a change in talker. Specifically, listeners who did not detect the change in the talker had a greater difference between conditions of lexical difficulty than listeners who noticed the change, or listeners who heard the same talker throughout. These results suggest that failures to detect changes are not limited to the visual domain and that processing at some level may be affected by changes in the environment. Furthermore, these results have implications for models of spoken word recognition.

1pSC7. Formant continuity and auditory scene analysis: The effect of vowel formant manipulations on the perception of synthetic nasal consonants.
Susan M. Harding and Georg F. Meyer (Mackay Inst. of Commun. and Neurosci., Keele Univ., Keele, Staffs ST5 5BG, UK, s.m.harding@cns.keele.ac.uk)

The perception of a synthesized nasal /m/ changes to /n/ as the frequency of the second formant of a preceding vowel is increased when there are no transitions between the vowel and nasal; but /m/ is heard consistently when 20-ms transitions are introduced. A possible explanation is provided by auditory scene analysis: Formants of the vowel and nasal that are contiguous and close in frequency may be grouped using principles of similarity and good continuity into a single perceptual stream. Further experiments found a robust change in percept for prototype /n/ as well as /m/. Transitions from the vowel that conflict with the formant structure of the consonant also cause a similar change in percept for both nasal prototypes. The proximity of the second formant of the vowel to a formant in the nasal prototype is therefore unnecessary for the change in percept to occur; the presence of the vowel formant near to one of two target frequencies at the boundary with the nasal seems to be sufficient to determine the nasal percept and takes precedence over the structure of the nasal prototype. Thus, these results do not show strong evidence for auditory scene analysis applied to formants.

1pSC8. Perceptual experiment on Korean number production.
Byunggon Yang (Dongeui Univ., 24 Kayadong, Pusanjingu, Pusan 614-714, Republic of Korea, bgyang@hyomin.dongeui.ac.kr)

Acoustic parameters of the nine Korean numbers were analyzed by PRAAT, a speech analysis software, and synthesized by SENSYNPPC, a Klatt formant synthesizer. The overall intensity, pitch, and formant values of the numbers were modified dynamically by a step of 1 dB, 1 Hz, and 2.5%, respectively. This study explored the sensitivity of listeners to changes in the three acoustic parameters. Twelve male and female subjects listened to 390 pairs of synthesized numbers and judged whether the given pair sounded the same or different. Results showed that subjects perceived the same sound quality within the range of 6.6 dB of intensity variation, 10.5 Hz of pitch variation, and 5.9% of the first three formant variations. The male and female groups showed almost the same perceptual ranges. Also, an asymmetrical structure of high and low boundary was observed. The ranges may be applicable to the development of a speaker identification system while the method of synthesis modification may apply to its evaluation data. [This work was supported by a grant from the Interdisciplinary Research Program (Contract No. 1999-2-302-106-5) of the KOSEF.]

1pSC9. Discrimination and labeling differences in audiovisual spech perception.
Nicole L. Marrone and Arlene E. Carney (Dept. of Commun. Disord., Univ. of Minnesota, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

In an earlier report, Clement *et al.* [J. Acoust. Soc. Am. **107**, 2887 (2000)] showed that listeners' confidence ratings of the identity of audiovisual syllables produced with conflicting auditory and visual cues were typically lower than those for audiovisually congruent syllables. This occurred even when both types of syllables were labeled consistently as the same phoneme. Confidence ratings varied across listeners, suggesting a graded nature of auditory–visual integration even across listeners who exhibit visual bias. In the current experiment, listeners were asked to identify audiovisual tokens from two talkers who elicit a high proportion of visually biased responses [Carney *et al.*, J. Acoust. Soc. Am. **106**, 2270 (1999)] and to rate their confidence in their labeling judgment. Only listeners who exhibited strong visual bias were tested in a discrimination task in which a true audiovisual /di/ was a standard. Comparison stimuli were visual /gi/-auditory /bi/, visual /bi/-auditory /gi/, and true audiovisual /bi/, /di/, and /gi/. Listeners were able to discriminate the visually biased and true /di/ stimuli while labeling them both as alveolar tokens. Results support the notion of a graded perception of bimodal stimuli. [Research supported by NIDCD.]

1pSC10. Sinewave speech/ nonspeech perception: An fMRI study.
Einat Liebenthal, Jeffrey R. Binder (Neurology Dept., Medical College of Wisconsin, 8701 Watertown Plank Rd., Milwaukee, WI 53226, einatl@mcw.edu), Rebecca L. Piorkowski, and Robert E. Remez (Barnard College, New York, NY 10027)

Sinewave speech replicas follow the center frequency and amplitude of vocal tract resonances of natural speech but lack the attributes of vocal chord vibrations. They can be perceived as speech or nonspeech depending on expectation (Remez *et al.*, 1981). Brain areas activated by sinewave replicas when perceived as nonspeech or as speech were compared with functional magnetic resonance imaging. In an auditory task, 30 nave subjects determined whether isolated second-tone formants were included in a following three-tones sinewave complex. The second-tone formant was either aligned (phonetic stimulus) or temporally reversed relative to the other tones in the complex (auditory stimulus). Halfway through the auditory task, subjects were informed about the phonetic derivation of the stimuli and trained to recognize the original speech. In a subsequent phonetic task, subjects identified sinewave stimuli containing the phoneme /p/. Following training, an increase in behavioral response time, concurrent with a decrease in left Heschl's gyrus activation was observed specifically with the phonetic stimuli. This suggests automatic

recognition of speech structure in the auditory cortex and interference with auditory processing of phonetic stimuli, acquired by training. In the phonetic task, the left inferior frontal gyrus was specifically activated implicating this area in phonetic processing.

1pSC11. Perceptual metathesis of obstruent clusters. Matthew J. Makashay (Dept. of Linguist., Ohio State Univ., 222 Oxley Hall, 1712 Neil Ave., Columbus, OH 43210, makashay@ling.ohio-state.edu)

This study examined acoustic and perceptual cues of obstruent clusters in order to test the hypothesis that metathesis can be a process that maintains identification of the consonants involved. In an auditory lexical decision task, there were effects of both optimality of cues, and lexical frequency of clusters. Nonword targets were created by metathesizing medial obstruent clusters in English words (e.g. [nækpɪ n] for *napkin*). For the clear listening level group, there was a slow rejection of nonword targets with optimal clusters that occur with higher frequency in the lexicon. For the speech reception threshold group, nonword targets with optimal clusters were more likely to be perceptually metathesized and realized as words than nonword targets with nonoptimal clusters were, presumably because subjects are more likely to hear both consonants in optimal clusters. Clusters with fricatives and stops were less likely to be perceptually metathesized than clusters containing only stops, since the continuity of manner features in a stop cluster hinders perception of consonant order. Whether some listeners are more adept than others in the ordering of temporal acoustic events will be investigated as well. [This material is based upon work supported under a National Science Foundation Graduate Fellowship.]

1pSC12. Flexibility of acoustic cue weighting in children's speech perception. Catherine Mayo, Alice Turk (Dept. of Theoretical & Appl. Linguist., Univ. of Edinburgh, Edinburgh EH8 9LL, UK), and Jocelynne Watson (Univ. of Edinburgh, Edinburgh EH9 1UW, UK)

Nittrouer and colleagues [Nittrouer, J. Phonetics **20**, 1–32 (1992); Nittrouer and Miller, J Acoust. Soc. Am. **101**, 2253–2265 (1997); Nittrouer *et al.*, Percept. Psychophys. **62** (2000)] have found that in identifying certain syllable contrasts, young children make more use of syllable–internal formant transitions (relative to other available acoustic cues) than do older children and adults. The evidence for this change in the degree to which listeners weight, or use, certain cues comes predominantly from studies of fricative contrasts (e.g., /sV/–/ʃV/, /sV/–/stV/, /Vs/–/Vʃ/). The current study tests the flexibility of children's weighting of acoustic cues by examining cue weighting across a wider range of phonetic contexts. In particular, this study attempts to determine whether children's focus of perceptual attention can be led away from transitions in contexts where such cues are relatively less salient. Additionally, the study tests children's ability to identify phonemes in an extreme situation, in the complete absence of transitional information. [Work supported by Wellcome Trust.]

1pSC13. Backward masking in speech perception by children and adults. Joan Sussman and Elizabeth Laczi (Dept. of CDS, Univ. at Buffalo, 105 Cary Hall, 3435 Main St., Buffalo, NY 14214, jsussman@acsu.buffalo.edu)

The current investigation measured discrimination and identification of tone-noise and [b]-[a] stimuli by adults, children with normally developing language, and children with language impairment. The children were aged 4–6 years. These tasks were chosen instead of the more traditional two alternative forced choice (2AFC) technique because it was believed possible that task difficulty influenced previous results (e.g., Wright *et al.*, 1997). The stimuli in the current set of experiments had longer tones (40 ms) than prior studies but used higher masker levels (52 dB spectrum level). However, results showed that all the participants discriminated the

stimuli containing the tone or [b] signal significantly above chance, contrary to previous findings. More difficulty was found in the identification task, similar to prior backward masking results. [Supported by an Individual Development Award, NYS/UUP.]

1pSC14. Overt visual attention in spoken sentence perception. Charissa R. Lansing (Dept. of Speech and Hearing Sci., Univ. of Illinois at Urban-Champaign, 901 S. Sixth St., Champaign, IL 61820, crl@uiuc.edu) and George W. McConkie (Beckman Inst., Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801)

People's eye movements and performance accuracy were recorded as they attempted to understand sentences spoken by two talkers under two conditions: vision only and vision plus low-intensity sound. Percent word-correct scores were higher for the vision-plus-sound than for the vision-only presentation and for the male compared to the female talker. Eye movement records showed a tendency to gaze at the talker's eyes when the talker was not speaking, but to shift the gaze to the mouth and make long eye fixations when the talker was speaking, particularly under vision-only conditions and for the female talker. In a task requiring verbatim word identification, people with average speech-reading proficiency direct their gaze to the talker's mouth most of the time during the talker's speech production, contrary to the finding of Vatikiotis-Bateson, Eigsti, Yano, and Munhall (1998), and they produce very long eye fixations. For these people, the gaze is drawn to the mouth, not by facial motion alone, but also on some other basis that is assumed to be prior knowledge of the location of critical visual cues, with an accompanying suppression of saccadic activity. [Work supported by NIH-NIDCD Grant DC02250.]

1pSC15. Limits of sentence identification in gated and continuous noise. Peggy B. Nelson, Su-Hyun Jin, and Arlene Earley Carney (Dept. of Commun. Disord., Univ. of Minnesota, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

Listeners with normal hearing sensitivity are able to take advantage of temporal dips in fluctuating noise. Their word identification in gated noise is better than identification in continuous noise. The limits of this ability are not well understood. Young adult listeners with normal hearing sensitivity were tested for their understanding of speech in gated and continuous noise at a variety of signal-to-noise ratios (SNRs) and gate frequencies from 2–32 Hz. Stimuli were digitized IEEE and CID sentences spoken by 10 talkers (5 male and 5 female). Pseudorandom noise was generated with the same long-term spectrum as the speech. Sentences were presented at an overall level of 65 dB SPL, with noise at +16-, +8-, 0-, −8- and −16-dB SNR. Listeners' responses were scored for the number of keywords correct. Results showed that at −16-dB SNR of continuous noise, listeners were unable to repeat any keywords. Gate frequencies of 4 and 8 Hz resulted in the greatest keyword identification. Performance decreased at the slowest gate frequency for IEEE sentences and at fastest frequencies for all sentences. Results will be discussed in terms of listeners with hearing loss and cochlear implants. [Work supported by NIDCD.]

1pSC16. Spatial distribution of early reflections and speech intelligibility. Yang-Ki Oh (Mokpo Natl. Univ., Mokpo, Chonnam 534-729, South Korea), Dae-Up Jeong (Chonbuk Natl. Univ., Jeonju, Chonbuk 561-756, South Korea), Se-Jin Doo (Dong-Ah Broadcasting College, Ansung, Kyonggi, South Korea), Hee-Won Lee (Seoul Natl. Univ. of Technol., Seoul, South Korea), Chul-Min Choi, Lai-Hoon Kim (Seoul Natl. Univ., Seoul, South Korea), and Il-Doo Ko (Seoul Natl. Univ. of Technol., Seoul, South Korea)

The strong early reflections and short delay times have been known to improve the intelligibility of speech heard in rooms. D50 and C80, the most frequently used physical parameters, were developed taking this fact into consideration. However, these monaural parameters have limited applications for the practical design of rooms because of their lack of spatial

information. The present work investigates how temporal changes in three-dimensional distribution of early reflections influence speech intelligibility in rooms. A new measurement method, using a five microphone array and an omnidirectional source setup, is employed, and a series of post-processing procedures are involved, for getting different early reflections in their spatial distributions. The changes were made for the impulse responses obtained through a five microphone array in the arrival times of early reflections from all, and the horizontal and vertical directions, respectively. Anechoic samples of the Korean language were convolved binaurally with the reproduced impulses by applying a head-related transfer function. A series of speech intelligibility tests, conducted for 22 university students, found that the percentage of correct responses significantly deteriorated by increasing delay times of early reflections from the vertical direction. The result suggests that vertical components of early reflections play a significant role in improving speech intelligibility. [Work supported by Korean Research Foundation Grant KRF-1999-1-310-004-3.]

1pSC17. Consonants and vowels discriminated differently even when acoustically matched.

A. Min Kang (Haskins Labs., 270 Crown St., New Haven, CT 06511 and Yale Univ., New Haven, CT 06520, min.kang@yale.edu)

Vowels are reportedly discriminated differently from consonants, but there have typically been large between-class acoustic differences. Discrimination still differed when acoustic differences were reduced by removing the mostly vocalic center portion of CVCs [silent center (SC)] [A. M. Kang and D. H. Whalen, J. Acoust. Soc. Am. **107**, 2855–2856 (2000)]. The present study compared consonant and vowel identification and discrimination of synthetic CVCs varying in equal-sized $F2$ steps along /b-d/ and /ɛ-ʌ/ continua (full syllables), and in truncated syllables corresponding to the initial 60 ms of the previously examined SC syllables. To lower listener uncertainty, only consonant, or only vowel, information was varied within a test block. Consonant discrimination for full syllables was much higher than in the earlier SC experiment; it was slightly higher for the truncated stimuli than for the full. Vowel discrimination was much higher than consonant, near ceiling for both full and truncated stimuli. Thus, even when acoustic steps are equalized and the speech presented (in the truncated stimuli) is limited to the syllable portion that contains most of the constant information, vowels remain better discriminated than consonants. This indicates a true difference processing of the two phonetic classes, even when the acoustics are well matched. [Work supported by NIH.]

1pSC18. Perception of Cantonese Parkinsonian speech.

Patrick C. M. Wong, Randy L. Diehl (Dept. of Psych., Univ. of Texas, Austin, TX 78712), Shu Leong Ho, Leonard S. W. Li (Univ. of Hong Kong, Hong Kong, PROC), and Kin Lun Tsang (Queen Mary Hospital, Hong Kong, PROC)

The current study is a continuation of our previous case study investigating the effect of reduced pitch range in Parkinsonian speech on a tone language [P. C. M. Wong and R. L. Diehl, J. Acoust. Soc. Am. **105**, 1246(A) (1999)]. In the first experiment, listeners were asked to identify the last word of semantically neutral sentences produced by Cantonese-speaking Parkinson's disease (PD) patients, normal speakers, and a resynthesized version of PD speech with expanded pitch range. Identification of normal and PD speech did not differ, perhaps due to the insignificant difference in pitch range between the two types of speech. However, listeners were better at identifying the resynthesized PD speech which contained a larger pitch range than the original PD speech. This latter result supports the theory of context-target pitch distance proposed by Wong and Diehl which states that lexical tone perception relies on a sufficiently large pitch distance between the context and target of an utterance [J. Acoust. Soc. Am. **104**, 1834(A) (1998)]. In the second experiment, subjects were asked to identify the intended intonation (angry, happy, neutral, and question) of sentences produced by normal and PD speakers. Performance was better for normal speech. [Work supported by NIDCD.]

1pSC19. Do listeners to speech perceive gestures? Evidence from choice and simple response time tasks.

Carol A. Fowler, Julie M. Brown, Laura Sabadini-Grant (Haskins Labs. and Dept. of Psych., Univ. of Connecticut, 406 Babbidge Rd., Unit 1020, Storrs, CT 06269, fowler@tom.haskins.yale.edu), and Jeffrey Weihing (Haskins Labs., New Haven, CT 06511-6695)

According to the motor and direct realist theories, listeners perceive speech gestures. The following experiments test this claim. Experiments 1 and 2 replicate the findings of Porter and Castellanos [J. Acoust. Soc. Am. **67**, 1349–1356 (1980)]. Participants shadowed vowel-consonant-vowels (VCVs) produced by a model. Responses were timed. The difference between response times (RT) in simple and choice speech shadowing tasks (26 ms) is shorter than the canonical choice/simple RT difference [100–150 ms, Luce (Oxford, New York, 1986)]. This is interpreted as supporting Porter and Castellanos, in that when the task is to shadow speech, the element of choice is considerably reduced as the listener receives instructions for her response from the speech sounds she perceives. In experiment 3, the timing of gestures of the models' speech was manipulated by extending the voice onset time (VOT) of the models' production of voiceless stops in half of the speakers VCVs. VOTs of participants shadowed responses were measured. Our findings suggest that listeners' productions of phonemes can be influenced by their perception of the timing of the models' gestures in speech shadowing tasks. This provides additional support for the interpretation that participants' shadowing responses are guided by their preception of the models' gestures.

1pSC20. Bandpass filtered faces and audiovisual speech perception.

Kevin Munhall (Dept. of Psych. and Otolarngol., Queen's Univ., Kingston, Canada), Christian Kroos, and Eric Vatikiotis-Bateson (ATR Intl.—Information Sci. Div., Kyoto, Japan)

The visual system processes images in terms of spatial frequency-tuned channels. However, it is not clear how complex object and motion processing are influenced by this early visual processing. In two studies this question was explored in audiovisual speech perception. Subjects were presented with spatial frequency filtered images of the moving face during a speech in noise task. A wavelet procedure was used to create five bandpass filtered stimulus sets. The CID Everyday sentences were presented with a multivoice babble noise signal and key word identification accuracy was scored. Performance varied across the filter bands with peak accuracy being observed for the images containing spatial frequencies spanning 7–14 cycles/face. Accuracy for higher and lower spatial frequency bands was found to be lower. When viewing distance was manipulated no change in the overall shape or peak in the key word accuracy function was observed. However, at the longest viewing distance the performance in the highest spatial frequency band decreased markedly. The results will be discussed in terms of visual information processing constraints on audiovisual integration.

1pSC21. Neighborhood effects in Japanese word recognition.

Kiyoko Yoneyama and Keith Johnson (Dept. of Linguist., Ohio State Univ., 222 Oxley Hall, 1712 Neil Ave., Columbus, OH 43210, yoneyama@ling.ohio-state.edu)

This paper reports on the results of a naming experiment that investigated lexical neighborhood effects in Japanese word recognition. A naming experiment was conducted with 28 Japanese adult listeners. Each participant responded to 700 words that had varying neighborhood density (in terms of Greenberg–Jenkins' phoneme substitution, deletion, and insertion rules). The lexicon used for this calculation consisted of only nouns from the NTT Japanese psycholinguistic database [Amano and Kondo (1999)]. A preliminary regression analysis showed that such neighborhood density was negatively correlated with naming reacting time. The words with higher neighborhood density were responded to faster than those with lower neighborhood density. We plan to report further analyses that (1)

include prosodic information as another dimension of the neighborhood calculation in order to reflect the finding that prosodic information has a vital role in Japanese word recognition, and that (2) calculate neighborhood density based on the auditory properties of the words in the lexicon. Neighborhood density is measured by comparing similarity of audio files in the Amano and Kondo database. Altogether, we plan to discuss which level of lexical representation (a lower-level acoustic-auditory representation or a higher-level abstract phonemic representation) is used to calculate phonological similarity within the lexicon.

1pSC22. Experimental paradigm for examining the formation of complex auditory categories. Andrew J. Lotto (Dept. of Psych., Washington State Univ., Pullman, WA 99164) and Lori L. Holt (Dept. of Psych., Carnegie Mellon Univ., Pittsburgh, PA 15213)

Phonetic perception is at its essence a categorization task involving multiple imperfectly valid cues to category membership. To examine the general cognitive processes underlying the formation and utilization of complex auditory categories such as speech, we have created a set of stimuli crafted from 300-m bursts of white noise. These stimuli vary in onset duration (or attack) and the center frequencies of two spectral gaps. Categories of varying complexity can be created from this stimulus set and input distributions can be strictly controlled. In addition, category formation can be observed ''online'' in this microgenetic design as subjects make responses (identification and discrimination) while learning the categories. In one such task, subjects were trained with categories defined by boundary values across three attributes. After 10 h of training, the subjects demonstrated extensive learning despite the fact that the categories suffered from ''lack of invariance;'' that is, no cue could be relied on exclusively to determine category membership. This demonstrates the viability of this paradigm for examining complex categorization. Implications for speech categorization and phonetic acquisition will be discussed. [Work supported by NSF.]

1pSC23. The role of spectral contrast in the perception of stop consonants following vowels and their spectral complements. Jeffry A. Coady and Keith R. Kluender (Dept. of Psych., Univ. of Wisconsin, 1202 W. Johnson St., Madison, WI 53706)

The ability of listeners to recover speech information, despite dramatic articulatory and acoustic assimilation between adjacent speech sounds is remarkable and central to understanding perception of connected speech. In recent years, studies have revealed that, to some extent, auditory processes of spectral contrast compensate for assimilative effects of coarticulation. In the present studies, series of CV syllables varying acoustically in $F2$-onset frequency and perceptually from [ba] to [da] were identified either following front (e.g., [i], [e]) and back (e.g., [u], [o]) vowels or following complementary nonspeech spectra. These nonspeech stimuli were harmonic spectra in which the amplitudes of individual harmonics were the inverse of their amplitudes for those same harmonics in the vowels. A prediction of a spectral contrast account is that these complementary spectra should affect perception of following sounds in a manner complementary to that for the vowels from which they were modeled. This predicted effect was obtained. Nevertheless, one may expect that not all perceptual accommodation of coarticulation is explained by spectral contrast. In addition to these initial results, findings from experiments in which stimuli and task variables are manipulated to assess the influence of other processes will be reported. [Work supported by NIDCD DC04072.]

1pSC24. Effects of modulation and laterality on the perception of time-varying sinusoidal sentences in noise. Melanie M. Richter and Thomas D. Carrell (Univ. of Nebraska, 318 Barkley Ctr., Lincoln, NE 68583)

Researchers have shown amplitude modulation significantly increases the intelligibility of time-varying sinusoidal (TVS) sentences [T. Carrell and J. Opie, Percept. Psychophys. **52**, 437–445 (1992); J. Barker and M. Cooke, Speech Commun. **27**, 159–174 (1999)] and ameliorates the effect of multispeaker babble on intelligibility at moderate signal-to-noise ratios [T. Carrell, J. Acoust. Soc. Am. **93**, 2327 (1993)]. These effects have been attributed to the ability of amplitude comodulation to create auditory objects. However, other characteristics of acoustic signals cause components to group together, the classic example being sound location in the cocktail party effect [C. Cherry, J. Acoust. Soc. Am. **25**, 437–445 (1953)]. In natural environments, acoustic cues work together to assemble and segregate auditory objects for auditory scene analysis [A. Bregman, *Auditory Scene Analysis* (1990)]. The present study examined the interaction of amplitude modulation and laterality on the intelligibility of TVS sentences in noise. Forty females listened to TVS stimuli in a $2\times2\times2$ factorial design experiment: modulation (100 Hz versus unmodulated) by laterality (diotic versus dichotic) by noise (S/N=$+5$ dB vs -15 dB). Modulation and noise demonstrated large main effects in the expected direction, but effects of laterality were minimal.

1pSC25. Infant-directed speech helps adults separate different streams of speech. Rochelle S. Newman and Tammy Weppelman (Dept. of Psych., Univ. of Iowa, Iowa City, IA 52242)

Infants, like adults, are frequently faced with multiple people speaking at one time. What cues help them separate different voices? One possible cue is infant-directed speech (IDS). Speech directed toward infants differs from adult-directed speech (ADS) in a number of ways. IDS typically involves higher pitch, wider pitch excursions, longer durations, and may involve greater amplitude variability [Aslin (1993); Fernald and Simon (1984); Fernald *et al.* (1989)]. Many of these same features have been shown to be important for adult musical streaming. Furthermore, IDS may serve to make the caretakers voice more distinct from background speech (which is likely to be adult directed). This suggests that it might be easier to separate IDS from ADS than it is to separate two ADS samples. To test this, adult listeners were asked to shadow a target voice while a distractor voice spoke in the background. Performance when one of the voices was infant directed was compared with that when both voices were adult directed. We found a significant effect of speech register on stream segregation, such that listeners found it much easier to follow the target voice when it (but not the distractor) spoke in an infant-directed manner. [Work supported by NSF.]

1pSC26. The basic units of rate normalization. Rochelle S. Newman (Dept. of Psych., E11 Seashore Hall, Univ. of Iowa, Iowa City, IA 52242)

Individuals vary their speaking rate, and listeners use the duration of adjacent segments to adjust for these changes [Miller and Liberman (1979); Newman and Sawusch (1996)]. However, it has not been clear what these ''segments'' actually are. We examined whether two-phoneme sequences would produce two separate rate normalization effects (one for each phoneme), or only a single effect. We created series such as ''shkib-chkib'' and ''shwib-chwib'' and examined effects of the /k/ or /w/ and /I/ durations on perception of the initial contrast. Altering the /k/ duration resulted in a rate normalization effect different from that caused by varying the duration of the sequence ''kib'' as a whole. This suggests that /k/ and the vowel are treated as separate units of speech. However, we did not

find this pattern for the /w/. For sequences such as ''kih,'' which have obvious acoustic boundaries, each phoneme has a separate rate normalization effect. For sequences such as ''wih,'' without such obvious cues, the sequence is treated as a single (larger) segment by the rate normalization process. This suggests that rate normalization is not based on phonemes, but on segments with obvious acoustic boundaries, emphasizing the role of basic auditory processing in speech recognition.

1pSC27. Identification of multiple consonants in binaural presentation. Bom Jun Kwon (House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057)

How well do we process multiple speech sounds presented concurrently? Consonant identification was tested with normal-hearing listeners using 13 voiced consonants, which were recorded by 10 speakers and had similar durations of murmur or voice bar, in syllable-initial position. Listeners were presented with multiple (up to four) stimuli with a time delay between the stimuli via headphone either diotically or dichotically (to simulate spatial separation between sound sources). Results showed that while the identification scores decreased as the delay decreased, performance was still quite robust even with the zero delay. The scores for diotic presentation were lower than those for dichotic presentation, but the difference was not overwhelming compared to the decrease in the score due to the delay. The data suggest that (i) listeners are capable of processing and identifying multiple consonants in concurrent presentation, and (ii) the effect of onset differences is more substantial on identification of multiple sounds than binaural separation. In other words, the ''cocktail party'' effect can occur even when there is no spatial separation between sound sources. A possible mechanism of identification of multiple consonants, when the cues to segregate each stimulus is highly limited, will also be discussed. [Work supported by NIDCD and Quota Scholarship.]

1pSC28. Intelligibility of frequency-shifted speech. Jack M. Scott, Peter F. Assmann (School of Human Development, Univ. of Texas–Dallas, P.O. Box 830688, Richardson, TX 75083), and Terrance M. Nearey (Univ. of Alberta, Edmonton, AB T6G 2E7, Canada)

A significant fact about speech perception is that intelligibility can be preserved when the formant pattern is shifted up or down along the frequency scale. To study the relationship between fundamental frequency ($F0$) and spectrum envelope shifts, we used a high-quality vocoder (''Straight'') to process a set of 280 English sentences from the Hearing in Noise Test. Upward shifts in spectrum envelope (shift factors of 1, 2, 3, or 4) were combined with upward shifts in $F0$ (shift factors of 1, 2, or 4). Word recognition accuracy dropped from near perfect for a spectrum envelope shift factor of 1 (unshifted speech) to 77%, 24%, and 7% for shift factors of 2, 3, and 4, respectively. $F0$ shifts had a small but significant effect on identification. For a spectrum envelope scale factor of 1, increasing $F0$ by two octaves (shift factor of 4) resulted in a 16% decline. However, for spectrum envelope shift factors of 2 or greater, increasing $F0$ led to an improvement in performance. The results are consistent with the hypothesis that learned relationships between $F0$ and spectrum envelope affect the intelligibility of frequency-shifted speech.

1pSC29. English-speaking listeners' sensitivity to allophone appropriateness in real and nonword speech perception. Amanda K. Jones (Dept. of Linguist., UCLA, 405 Hilgard Ave., Los Angeles, CA 90025, amandajo@ucla.edu)

The question of whether adult listeners are capable of evaluating speech for the appropriateness of its allophonic content has received relatively little attention. Whalen *et al.* (1997) showed that English-speaking listeners preferred context-appropriate allophones when evaluating real words. However, for nonwords they found that listeners tended to prefer one particular allophone, regardless of its contextual appropriateness. In the current study, listeners provided naturalness ratings for pronunciations of both words and nonwords containing either an appropriate or an inappropriate allophone of the phoneme /t/ in initial /st/ clusters. The words and nonwords were divided into four categories determined by pretesting: highly familiar real words (e.g., ''start''), relatively unfamiliar real words (''stave''), highly wordlike nonwords (''storch''), and less wordlike nonwords (''stimf''). Results showed that listeners consistently preferred pronunciations that contained an appropriate allophone of /t/, regardless of stimuli category, e.g., real word vs nonword. This result contradicts Whalen *et al.*'s suggestion that ''allophonic contrast . . . may require contact with the lexicon in order to affect production and perception.'' Instead, the results indicate that English-speaking listeners have abstract concepts of allophonic appropriateness that they can apply to both existing lexical items and nonwords when evaluating these items for the naturalness of their pronunciations.

1pSC30. Automatic gender identification. Alireza Afshordi Dibazar, Jim-Shih Liaw, and Theodore W. Berger (Dept. of Biomed. Eng., Univ. of Southern California, OHE-500, Los Angeles, CA 90089-1451)

In this paper, a new automatic gender identification method is proposed as a part of an Automatic Speaker Recognition (ASR) system. The short time raw speech signal (180–450 ms) was filtered by the nine-order Butterworth low pass filter and decomposed to different frequency bands by the wavelet filter bank analyzer. The energy of the 120 sub-bands was used as a feature vector and applied to a standard classifier. This classifier was trained by the gradient descent method with 1594 utterances spoken by various males and females. The system was tested with different 2542 utterances giving 99.2% correct classification rate. The high performance and simplicity of implementation are the characteristics of this system in comparison to the other methods. [Work supported by DARPA.]

1pSC31. Efficient automatic recognition of spoken digit strings. Douglas O'Shaughnessy and Hesham Tolba (INRS-Telecommunications, 900 de la Gauchetiere west, P.O. Box 644, Montreal, PQ H5A 1C6, Canada)

Automatic recognition of spoken digit sequences (such as credit card numbers) is now feasible even in speaker-independent applications over the telephone. However, all recognition tasks have lower performance in noisy conditions. If significant limitations are also imposed on the computational resources used for recognition, then robust speech recognition is still a significant challenge, even for a simple digit vocabulary. Since recognition of continuously spoken digits over telephone links is a very practical application, such recognition was investigated here under different conditions. Traditional hidden Markov model approaches with cepstral analysis were not used, because they are computationally intensive and have not always worked well under adverse acoustic conditions. Simpler spectral analysis was used, combined with a segmental approach. The analysis focuses on locations of spectral peaks, similar to formant tracking, but without the need to estimate peaks for all time frames. The limited nature of the vocabulary (i.e., ten digits) allows this simpler approach. High recognition accuracy is maintained despite being very efficient in both memory and computation. Recent progress will be reported.

Session 1pSP

Signal Processing in Acoustics: Nondestructive Testing: Signal Processing

David H. Chambers, Chair

Electrical Engineering, Lawrence Livermore National Laboratory, L-154, P.O. Box 808, Livermore, California 94551-5508

Contributed Papers

2:30

1pSP1. Ultrasonic pulse compression system with new approaches of signal processing. Khan M. Mahmud and Ryoji Ohba (Div. of Appl. Phys., Grad. School of Eng., Hokkaido Univ., Sapporo 060-8628, Japan, mah@eng.hokudai.ac.jp)

A nonscanning method for ultrasonic flaw detection technique is described. The system employs an M-sequence modulated ultrasonic wave as the excitation signal. Irrespective of the number of faults to be detected, the data acquisition system predeterminedly consists of a transmitter and a few receivers to be fixed at convenient locations. The cross-correlation function (CCF) between the original M sequence and the demodulated received sequence indicates the presence of faults in the form of sharp peaks. In order to clearly detect the expected peaks corresponding to the faults in the CCF, new signal processing techniques are proposed. The novel design approach of inverse filter and synchronous moving average are found to be very effective in eliminating the false peaks from the CCF while increasing the signal-to-noise ratio significantly. A way of using the CLEAN algorithm from a diverse field is proposed, which shows tremendous potential in ultrasonic NDT for the distinction of close peaks removing all sidelobes. Experimental results (on metal plate) demonstrate the capability of the system and confirm the feasibility of the approach under heavy ambient noise condition. An algorithm, only on the basis of travel time of the signal, to determine the exact location of a fault is described. [The authors would like to express their gratitude to the Ministry of Educaion, Japan, for the funds needed for this research.]

2:45

1pSP2. Space resolution improvement in pulse crack location. Vyacheslav V. Kazakov, Vitaly A. Zverev (Inst. of Appl. Phys., RAN, Nizhny Novgorod, Russia), and Alexander M. Sutin (Stevens Inst. of Technol., Hoboken, NJ 07030)

The new method of space resolution increasing in pulse location method was suggested by Professor Zverev is applied for space resolution improvement of ultrasonic pulse location technique. According to this method, the spectrum of the received signal is divided to the spectrum of the reference radiated signal and the complex inverse Fourier transformation is applied to the divided spectrum. This approach extends the frequency band used for analysis and highly increases the space resolution of crack location. The feasibility test was conducted with pulse ultrasonic device working with carried frequency about 3 MHz. The suggested processing was applied to the rectified signal so the frequency band of processing was restricted by 700 kHz. Even in this not optimal processing we observed the shortening of the signal reflected from a crack more than 20 times. The suggested technique was also used for spatial resolution improvement of nonlinear modulation crack imaging based on modulation of ultrasonic pulses by a crack vibration. [Work was partially supported by International Science Technical Center Grant No. 1369.]

3:00

1pSP3. Processing of ultrasonic measurements for nondestructive evaluation of critical optical components. Alan Meyer and James Candy (Lawrence Livermore Natl. Lab., P.O. Box 808, L-154, Livermore, CA 94551)

Real-time nondestructive evaluation of critical optics in high-energy, pulsed laser experiments is crucial from a safety as well as maintenance viewpoint. Fluence levels in short pulse, high-energy lasers can produce pits and cracks in the surfaces of the lasers' optical components. These flaws in the optical glass can adversely effect the production of the laser light or even result in a catastrophic failure of the optical component itself. Consequently, the identification and characterization of these flaws is critical. The production requirements of our target laser system mandate the automation of any inspection system. Consequently, it is necessary to develop an autonomous solution to process the reflected acoustic signals and, using this information, detect, localize, and characterize the optical flaws. In this paper we develop sophisticated signal and image processing techniques to automatically analyze laser-induced damage in optical glass. After preprocessing the raw channel measurement data from two orthogonal, narrow beamwidth, transducer arrays, a two-dimensional power image is created. A physics-based 2D matched filter is developed and used to detect and localize the flaws. An iterative solution is developed to sequentially search the resulting 2D image to extract and characterize the flaws.

3:15

1pSP4. Inversion methods for Lamb and Rayleigh wave surface holography. Kenbu Teramoto (Dept. of Mech. Eng., Saga Univ., Saga-shi 840-8502, Japan) and Hajime Yuasa (Akishima Lab., Mitsui Zosen, Inc., Tokyo, Japan)

Lamb and Rayleigh wave surface holographies have advantages for nondestructive testing of concrete and steel structures that is based on the use of generated surface acoustic waves that propagate over the concrete and steel and are reflected by external cracks. When the surface wave field is observed by a circular array of surface wave transducers, the proposed method that is called ''eigenfunction series expansion (EFSE)'' can reconstruct exact images of the scatterers within the array. This holographic imaging system has the following features: (1) An inversion via expansion in eigenfunction series produces much higher resolution than one which a back propagation produces. (2) An iso-azimuthal signal processing enables to design a first image reconstruction procedure. In this study, the computational process in EFSE and wave fields near the cracks are discussed and their physical meanings are investigated through FEM simulations. The surface wave transfer function by EFSE permits an effective regularization of the source field with cracks: this is possible because the transfer function includes several wave modes at the cracks and surface wave fields. In particular, it is clearly shown that the reconstructed image can be improved by adopting EFSE method that suppresses the spurious lobes.

2317   J. Acoust. Soc. Am., Vol. 109, No. 5, Pt. 2, May 2001

141st Meeting: Acoustical Society of America   **2317**

3:30

1pSP5. Application of subband adaptive filtering techniques to ultrasonic detection in multilayers. Jie Mao and MingXuan Li (Inst. of Acoust., Chinese Acad. of Sci., 17 Zhongguancun St., Beijing 100080, PROC, jiem@sohu.com)

The subband adaptive filtering method used in the evaluation of multi-layered steel–rubber debonding is discussed in the paper. Because the acoustic impedance of the steel differs far from that of the couplant water and the rubber, and the signal decay in the rubber is heavy, the energy of the signal reflected from the debonded rubber layers is very weak. The echoes caused by the steel/rubber interface almost mask the flaw echoes, which makes the extraction of the debonding echoes very difficult. In recent years, adaptive deconvolution [X. M. Jian, Chin. J. Acoust. **18**, 280–288 (1999)], wavelet inverse filtering, homomorphic filtering and echo cancellation techniques are adopted to extract the flaw echoes, but the signal reflected from the water/steel interface is ignored because of its large amplitude. A subband adaptive filtering method is discussed in the paper, where the subband decomposition is performed using a mutual wavelet packets decomposition on the criterion of maximizing the cross-correlation between the signals. The simulations on both synthetic and real signals are presented. [Work supported by National Natural Science Foundation of China.]

3:45

1pSP6. Acoustic scattering from laser damage spots. David H. Chambers, L. Peter Martin, and Graham Thomas (Lawrence Livermore Natl. Lab., P.O. Box 808 L-154, Livermore, CA 94551)

High energy laser systems sustain damage to critical optical components over their lifetime. This damage is manifested as localized pitting and cracking in the glass components. The damage pits have complicated morphology consisting of a hemispherical pit lined with crushed material, and surrounded by an outer zone with cracks. Monitoring of damage is necessary in some applications to prevent catastrophic failure of critical components. An acoustic system for monitoring optical damage spots was investigated to determine its applicability for the National Ignition Facility (NIF) at LLNL. Experiments and simulations of acoustic scattering from damage spots were performed to assess the ability to detect and size optical damage before failure becomes imminent. Pulses with a center frequency of 5 MHz were scattered from damage spots with nominal sizes of 0.5 mm to 7 mm. The amplitude of the measured return showed evidence of resonant scattering from the complicated morphology of the pits. This was confirmed with numerical simulations of scattering from simplified models of damage spots. This talk will emphasize the physics of scattering from damage spots and the comparison between experiments and numerical simulations.

4:00

1pSP7. A new acoustical method for inspection of industry piping. Ishtiaq Rasool Khan and Ryoji Ohba (Div. of Appl. Phys., Grad. School of Eng., Hokkaido Univ., Sapporo 060-8628, Japan)

A new acoustical method for inspection of the industry piping against development of corrosion and early stage cracks is presented. Impulse response of a defective pipe is obtained by launching an M-sequence at an end of the pipe, and cross-correlating it with the sequence received at the same end. The received sequence contains several replica of the M-sequence, reflected from various objects including the defects in the inner wall of the pipe. Therefore, the impulse response contains a peak corresponding to each reflection, the location of which determines the location of the reflecting object. Signal processing techniques are used to detect these very small peaks from heavy noise, and to distinguish the wanted peaks (due to defects) from unwanted peaks (due to other objects). A new technique is presented, which removes all kinds of unwanted peaks and other noise and significantly enhances the wanted peaks. This technique is proved much better than the classical noise removal techniques of moving average, subtraction and inverse filtering. A very simple experimental setup is presented and the results of successful detection of defects of different sizes in pipes of different dimensions are presented. [The presented research was financially supported by Japan Society for Promotion of Science.]

MONDAY AFTERNOON, 4 JUNE 2001

MONROE ROOM, 2:00 TO 4:45 P.M.

Session 1pUW

Underwater Acoustics: Matched Field Processing

Gregory J. Orris, Chair

Naval Research Laboratory, Code 7145, 4555 Overlook Avenue, SW, Washington, DC 20375-5000

Contributed Papers

2:00

1pUW1. Locating a sound source of unstable frequency by use of phase tracking. Travis L. Poole (MIT/WHOI, Joint Prog. in Oceanogr. and Oceanogr. Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543) and George V. Frisk (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

When the frequency of a cw source is known precisely, it is possible to relate the time rate-of-change of the measured phase to the time rate-of-change of the distance between source and receiver. However, when the source frequency is not known precisely, or when the source frequency is unstable, an unknown time dependence will remain after demodulation, causing the phase to appear to vary rapidly. Thus, the phase cannot be used to find the rate of separation of the source and receiver. By examining the difference between the phase measured by two independently moving receivers, it is possible to eliminate this unknown time dependence from the phase. In this talk, a source localization method based on this idea is presented and applied to data from a frequency-unstable source (near 20 Hz) in a shallow-water, low-frequency acoustics experiment. [Work supported by ONR and the WHOI Education Office.]

2:15

1pUW2. Matched-phase matched field processing: Searching for the phases. Gregory J. Orris and Michael Nicholas (U.S. Naval Res. Lab., Washington, DC 20375)

There have been several methods explored in the literature by which standard single-frequency matched-field techniques can be extended to multifrequency coherent algorithms. The foremost difficulty in implementing any coherent algorithm is the lack of reliable estimates of the relative phases of the processed frequency components. Thus many coherent techniques have side-stepped this issue by assuming an inter-frequency phase relationship that is either known *a priori*, or trivially zero. To address this

deficiency we have developed the matched-phase matched-field processor [Orris *et al.*, J. Acoust. Soc. Am. **107**, 2563–2575 (2000)]. The approach is to enlarge the search space from only the spatial dimensions to include the interfrequency phases. This processor has been applied to data from a source of known characteristics and was found to increase the spatial resolution while simultaneously increasing array performance significantly (received levels below −15 dB). We also present an application to a source of opportunity (i.e., a passing surface vessel) with unknown acoustic signature.

2:30

1pUW3. Information theory for matched field processing: preliminary results. John R. Buck (ECE Dept. and SMaST, Univ. of Massachusetts–Dartmouth, 285 Old Westport Rd., North Dartmouth, MA 02747)

Matched field processing (MFP) has traditionally been considered as an estimation problem in which the goal is to use the observed pressure field and an acoustic propagation model to estimate an unknown source location with the smallest possible variance. An alternative perspective is to divide the search region into cells whose size is dictated by logistical constraints, and attempt to assign the source to one of these cells. In this perspective, the goal of MFP is to minimize the probability of error in assigning the source to a grid cell, rather than minimizing the variance of the source location estimate. This perspective leads one to consider MFP as a communication problem in which the source transmits its location to the receiver array. Information theory provides techniques which can bound the amount of information available to the receiver in this problem. The amount of information observable has implications for the performance limits achievable by any MFP algorithm. Additionally, this perspective leads to methods of designing arrays for MFP which maximize the average information observed over the search space. [Work supported by ONR Young Investigator Program.]

2:45

1pUW4. Single snapshot array processing: The CLEAN algorithm applied to the pressure field. J. de Rosny, H. C. Song, and W. A. Kuperman (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0701)

The CLEAN algorithm, originally developed in radio astronomy [Astron. Astrophys. Suppl. **15**, 417–426 (1974)] for intensity images, has been modified to process complex pressure signals. The iterative method allows one to find the position and strengths of a distribution of sources which are assumed to be grouped in small regions of the sky. Recently, we have applied the CLEAN algorithm successfully to matched field processing where the ambiguity surface corresponds to an image. The original CLEAN algorithm, however, is applied only to the field intensity (magnitude). In this paper, we apply the CLEAN algorithm to the complex pressure field received by a vertical array to enhance the matched field processing. This imaging method appears to be useful when just one or a few snapshots are available so that a reliable cross-spectral density matrix cannot be obtained. Numerical simulations will be presented for the modified CLEAN algorithm that extracts a weak source masked by a stronger source.

3:00

1pUW5. Improvement in matched field processing by the CLEAN algorithm. H. C. Song, J. de Rosny, and W. A. Kuperman (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0238)

Matched field processing can be improved by the CLEAN algorithm [Astron. Astrophys. Suppl. **15**, 417–426 (1974)], introduced into radio astronomy for image restoration. The CLEAN algorithm is based on the *a priori* assumption that most images of radio sources are essentially blank sky with point-source components distributed in small regions. A simple iterative approach is employed to find the positions and strengths of these point sources. On convolving the point-source components with a ''clean'' beam, a final image known as the ''clean'' map is derived. In this paper,

we apply this approach to match field processing where the ambiguity surface corresponds to the astronomical map. The initial ambiguity surface called a ''dirty'' map is evolved to a clean map through the iteration process. Numerical simulations demonstrate that matched field processing combined with the CLEAN algorithm provides better performance especially when a weak source is hardly visible due to sidelobes generated by a much stronger source.

3:15

1pUW6. Source localization through knowledge of the scattered field near two air-filled spheres: Numerical modeling results. Paul A. Lepper, Gerald L. D'Spain, and William S. Hodgkiss (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704)

Knowledge of the complex nature of the underwater sound field measured near two scatterers that are in the far field of an acoustic source may provide valuable information on the location of the source. Traditional techniques of source localization such as time-of-arrival (phase) and amplitude differences measured by multiple receivers are complicated when placed within the near field of an acoustic scatterer due to effects such as diffraction and scattering. The use of the interference pattern field due to the presence of the scatterers is proposed to enhance traditional source localization techniques. Numerical results are presented using a 2-D finite-difference time-domain (FDTD) scheme for scattering from two fluid-loaded, thin-walled, air-filled elastic spheres for multiple source aspect angles and frequencies. Matched field processing techniques are used to quantify the source localization capabilities. [Work supported by ONR.]

3:30–3:45 Break

3:45

1pUW7. Array element localization for towed marine seismic arrays. Stan E. Dosso and Michael Riedel (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, sdosso@uvic.ca)

This work presents a new approach to array element localization (AEL) for the sensors of a towed marine seismic array based on regularized inversion of direct and bottom-reflected acoustic ray travel times picked from recorded seismic sections. Depth-sensor measurements at a number of points along the array are included as *a priori* estimates (with uncertainties) in the inversion. The smoothest array shape consistent with the acoustic data and prior estimates is determined by minimizing the array curvature or roughness. A smooth array shape is physically reasonable; in addition, minimizing array curvature provides an *a priori* model for the correlation between hydrophone positions that allows the estimation of both the offset and depth of hydrophones that record only one (or even no) acoustic arrival due to the shadowing effects of water-column refraction or reflection from arbitrary bathymetry. The AEL inversion is applied to a 102-sensor, 1.2-km towed array to correct receiver positions in the seismic velocity analysis of a seabed gas hydrate survey.

4:00

1pUW8. Matched field processing using a large aperture planar towed array. Jennifer Munro (MIT Lincoln Lab., 244 Wood St., Lexington, MA 02420), Arthur Baggeroer (MIT, Cambridge, MA 02139), James Ward, and Brian Tracey (MIT Lincoln Lab., Lexington, MA 02420)

Passive detection and localization in shallow water is complicated by high transmission loss and coherent multipath generated by bottom interaction. To compensate for this loss, one must achieve as much array gain as possible. The seismic community has developed large aperture arrays designed for low-frequency operation, which are on the order of 3–10 km by 0.5 km in size. This presentation explores the applicability of these arrays to passive detection and localization of underwater targets using matched field processing (MFP). This approach offers the advantages of a high number of sensors, large gain focusing, and three-dimensional localization through MFP when the propagation environment is well known. This talk presents simulated results of conventional MFP using planar

arrays in the presence of surface interferers. This presentation continues showing improvement through adaptive MFP. It then examines the problem of target motion. For these large arrays, robustness to target motion becomes a significant challenge. [This work was sponsored by the Defense Advanced Research Projects Agency under Air Force Contract No. F19628-00-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and not necessarily endorsed by the United States Air Force.]

4:15

1pUW9. Matched-field processing in shallow water in the presence of internal waves. Catherine Stamoulis (Dept. of Ocean Eng., MIT, Cambridge, MA 02139 and Naval. Res. Lab., Washington, DC 20375)

The performance of matched-field processing algorithms in shallow water, in the presence of soliton packets, has been investigated through analysis of data from the SWARM95 experiment. In particular, acoustic data collected with a vertical array and a moving up-sweep source have been analyzed, in the frequency range 275–350 Hz. The source was placed below the surface mixed layer and its distance from the array varied between 2.5 and 27 km. Source localization using different matched-field algorithms has been performed, in order to assess the robustness of these processors in the presence of environmental mismatch and their degradation as a function of range. The primary result of the analysis is that these processors failed to localize the source accurately, even at the shortest range (2.5 km). Although broadband processing and inclusion of corrections for array tilt improved the results, the errors in

source range (of the order of 1 km at short ranges and 5 to 10 km at long ranges) and depth remain significant. Alternative processing via mode filtering and subsequent matched-mode processing yielded more accurate results, particularly in cases were mode coupling was not an issue of concern.

4:30

1pUW10. Multitarget tracking using simulated annealing. David P. Knobles (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029), Tracianne Neilsen, and Robert A. Koch (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713)

It is common for passive detection systems to have multiple detections over a given time interval. The focus of this research is the simultaneous inversion for environmental parameters and the localization of multiple moving broadband sources in shallow water environments. Individual sources are isolated using track beam methods on array subapertures. The subaperture track beam spectra are then input as element data into an inversion approach that uses simulated annealing to estimate both environmental and source parameters. Environmental parameters include those associated with the seabed and source parameters include source depth, initial range, initial bearing, speed, and course. A key concept is that of space-time processing manifested in the cost function by coherent sums of product spectra over frequency, receiver pairs, and time sequence. Insight into the problem is obtained with simulated data by examining the coupling between individual parameters, including the coupling between source and environmental parameters. [Research supported by ONR.]

NOTE: Attendance at this session requires payment of an additional registration fee.

MONDAY EVENING, 4 JUNE 2001 STATE BALLROOM, 7:00 TO 10:30 P.M.

Session 1eID

Interdisciplinary: Tutorial Lecture on Demonstration Experiments: Videos and Audios for Teaching Acoustics

Thomas D. Rossing, Cochair
Physics Department, Northern Illinois University, DeKalb, Illinois 60115

Uwe J. Hansen, Cochair
Department of Physics, Indiana University, Terre Haute, Indiana 47809

Invited Paper

7:00

1eID1. Demonstration experiments, videos and audios for teaching acoustics. Thomas D. Rossing (Northern Illinois Univ., Dekalb, IL 60115), Uwe J. Hansen (Indiana Univ., Terre Haute, IN 47809), and a team of Chicago-area physics teachers

Since the time of Pythagorus, if not before, experiments have been used by physics teachers to illustrate acoustic phenomena and important acoustic principles. We owe much to Michael Faraday who established a rich tradition of lecture demonstrations at the Royal Institution in London. Lord Rayleigh, in addition to his well-known papers on the theory of sound, published a number of short notes which he entitled ''Acoustical Observations.'' A potpourri of acoustics experiments will be demonstrated by a team of college and high school teachers with instructions for doing these experiments with simple equipment. Reproduction of these demonstrations may be useful in a formal educational setting for teaching acoustics, and should be useful also in any number of informal venues— including ones not always thought of as ''teaching.'' Videos and audios demonstrating acoustics also will be previewed.