



## Free-field reciprocity calibration of laboratory standard (LS) microphones using a time selective technique

Rasmussen, Knud; Barrera Figueroa, Salvador

*Published in:*  
Journal of the Acoustical Society of America

*Publication date:*  
2006

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Rasmussen, K., & Barrera Figueroa, S. (2006). Free-field reciprocity calibration of laboratory standard (LS) microphones using a time selective technique. In *Journal of the Acoustical Society of America* (Vol. 120)

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

## Session 4aAA

## Architectural Acoustics: Measurement of Room Acoustics I

Fumiaki Satoh, Cochair

*Chiba Inst. of Technology, Tsudanuma 2-17-1 Narashino-shi, Chiba 275-0016, Japan*

Boaz Rafaely, Cochair

*Ben Gurion Univ., Electrical and Computer Engineering Dept., 84105, Beer Sheva, Israel*

Chair's Introduction—7:30

*Invited Papers*

7:35

**4aAA1. Warped-time-stretched pulse: An acoustic test signal robust against ambient noise.** Masanori Morise, Toshio Irino, Hideki Banno, and Hideki Kawahara (Wakayama Univ., 930, Sakaedani, Wakayama, 640-8510, Japan, s055068@sys.wakayama-u.ac.jp)

A new acoustic measurement signal that is a hybrid signal of time-stretched pulse (TSP), or lin-TSP, and logarithmic TSP (log-TSP) is proposed. The signal, referred to as warped-TSP [Morise *et al.*, IEICE Trans. Fundamentals, A, **J89-A**(1), 7–14 (2006)], has a single parameter to adjust for better measurements in accordance with ambient noise conditions. It also provides a means to eliminate harmonic distortions produced mainly by loudspeaker systems. In this lecture, the definition and features of the warped-TSP in comparison with the lin-TSP and log-TSP are introduced. The following were shown: (1) the relationship between the parameters, the amplitude frequency characteristics, and the effect on the harmonic distortion components; (2) a method to select the optimal parameters of the warped-TSP for a specific measuring environment; and (3) the experimental results for a series of impulse response measurements under different ambient noise conditions. Those results show that the proposed method outperformed the lin-TSP and log-TSP under all conditions in terms of SNR of the measured impulse response. [This research was supported partly by grants-in-aid for scientific research (15300061 and 15650032) and a grant from the Faculty of Systems Engineering at Wakayama University.]

7:55

**4aAA2. Simultaneous estimation of reverberation times and their uncertainties from room impulse responses using a single-measurement procedure.** Ning Xiang and Tomislav Jasa (Grad. Program in Architecture Acoust., and Dept. of Elec., Comput., and Systems Eng., Rensselaer Polyt. Inst, Troy, NY 12180)

Accurate measurements of reverberation times are of fundamental importance in room acoustics. A number of test procedures for characterizing acoustics in performing arts venues, quantifying acoustic properties of materials in chamber measurements, rely on experimental determination of reverberation times. In addition, decay-time estimation in acoustically coupled spaces has been found to be very demanding. Our recent work has demonstrated that model-based Bayesian approaches [Xiang *et al.*, J. Acoust. Soc. Am. **110**, 1415–1424 (2001); **113**, 2685–2697 (2003); **117**, 3705–3715 (2005)] can be very useful for such analysis in architectural acoustics measurements. This paper discusses the recent development of probabilistic tools for estimating both reverberation (decay) times and their uncertainties within Bayesian framework. This work shows that Bayesian probabilistic inference can be used as a useful tool for sound energy decay analysis in both single-space halls and coupled spaces. Bayesian decay analysis simultaneously provides architectural acousticians with reverberation times, diverse decay times, related derivations, and interdependencies to quantify uncertainties of the estimation from a single measurement of room impulse responses followed by Schroeder backward integrations.

8:15

**4aAA3. Permissible number of synchronous averaging times to obtain reverberation time from impulse response under time-variance conditions.** Fumiaki Satoh, Yukiteru Hayashi (Chiba Inst. of Technol., Tsudanuma 2-17-1, Narashino-shi, Chiba, 275-0016, Japan), Shinichi Sakamoto (Univ. of Tokyo, Meguro-ku, Tokyo, 153-8505, Japan), and Hideki Tachibana (Chiba Inst. of Technol., Narashino-shi, Chiba, 275-0016, Japan)

In the measurement of room impulse response, the synchronous averaging technique and such new methods as the MLS and the swept-sine methods are being widely used to improve the signal-to-noise ratio. In actual measurement conditions, however, the air in a room is continuously moving and the temperature is changing to some degree. The measured value of the reverberation time in such a room tends to be shorter at higher frequencies when applying the synchronous averaging. Therefore, the assumption of a time invariant has to be carefully considered, and, on this point, some research has been conducted to date. We also have reported various research results concerning the impulse response measurement under the time-variance conditions. In this paper, the permissible number of synchronous averaging times for reverberation measurement is studied through some field experiments. In each field, many

time impulse response measurements were taken between a fixed pair of sound source and receiving positions by the swept-sine method, without averaging. After the measurements, the characteristics and the extent of the time-variance under measuring were estimated by a short-term running cross-correlation function between each impulse response. The influence of the time variance on the synchronous averaging result was studied based on the estimated time variance.

8:35

**4aAA4. Selection of receiving positions suitable for evaluating acoustical parameters.** Taeko Akama, Hisaharu Suzuki, and Akira Omoto (Omoto Lab., Dept. of Acoust. Design, Faculty of Design, Kyushu Univ., Shiobaru 4-9-1, Minami, Fukuoka 811-8540, Japan)

Many physical parameters show characteristics of large sound fields such as concert halls. Some of them are adopted in the Annex of ISO 3382. That definition is clearly provided in ISO. However, practical measurement methods for them remain obscure. Our research is intended to examine an effective selection method of receiving positions based on the distribution of acoustical parameters in a real field. For that purpose, impulse responses are measured at more than 1400 seat positions to elucidate the distribution of acoustical parameters in an existing concert hall. The acoustical parameters, which are reverberation time, early decay time, clarity, and center time at each seat, are then calculated for 500-Hz, 1-kHz, and 2-kHz octave bands. The distributions of reverberation time are quite even at all seats. However, the distributions of other parameters show symmetrical patterns at 500 Hz. At 1 and 2 kHz frequencies, the distributions show asymmetrical patterns in this hall. Based on the results obtained in this study, an effective method to select the receiving position can be proposed.

8:55

**4aAA5. Hybrid measurement method in room acoustics using dodecahedron speakers and a subwoofer.** Hideo Miyazaki (Ctr. for Adv. Sound Technologies, Yamaha Corp., 203 Matsunokijima, Iwata, Shizuoka 438-0192, miyazaki@beat.yamaha.co.jp)

A dodecahedron speaker is usually utilized for measurement in room acoustics under the hypothesis of omni directional point source. But generally speakers used for a dodecahedron speaker cannot playback low-frequency sound such as under 100 Hz, which is important especially for auralization, while the one constructed of units with large diameter to support low-frequency sounds cannot be considered as an omni-directional speaker in high frequencies. To meet these requirements, a hybrid system combining a dodecahedron speaker and a subwoofer has been developed and actually used for measurements of impulse responses in acoustical design of concert halls. The summary of this method will be presented. The feasibility of this method will be also discussed while evaluating the measurement results in concert halls by changing measurement conditions such as speaker locations and comparing these results with those of conventional methods.

9:15

**4aAA6. The perception of apparent source width and its dependence on frequency and loudness.** Ingo B. Witew and Johannes A. Buechler (Inst. of Tech. Acoust., RWTH Aachen Univ., Templergraben 55, 52066 Aachen, Germany)

While it is widely accepted that apparent source width (ASW) is an important factor in characterizing the acoustics of a concert hall, there is still a lively discussion on how to refine the physical measures for ASW. A lot of experience has been gathered with interaural-cross-correlation and lateral-sound-incidence measures during the last years. As a result it was learned that different frequencies contribute differently to the perception of ASW and that the level of a sound also influences the perception of the apparent width of a source. With many technical measures having an influence on the perceptual aspect of ASW, the design of psychometric experiments becomes challenging as it is desirable to avoid the interaction of different objective parameters. In the experiments for the study presented, the perception of ASW is investigated for frequencies ranging from 100 Hz to 12.5 kHz at different levels of loudness. It is shown how the frequency and the loudness of a sound influence the perception of ASW.

9:35

**4aAA7. Sound source with adjustable directivity.** Gottfried K. Behler (Inst. fuer Technische Akustik, RWTH Aachen Univ., D-52056 Aachen, Germany)

Omni-directional sound sources are used to measure room-acoustical parameters in accordance with ISO 3382. To record a detailed room impulse response (RIR) with the aim of auralization, an extended frequency range is required that is not covered by the often-used building acoustics sound sources. To obtain this target, a loudspeaker with dedicated sources for low, mid, and high frequencies was designed, providing a smooth omni-directionality up to 6 kHz and a usable frequency range from 40 Hz up to 20 kHz. However, a realistic auralization of sources like musical instruments is not possible with an omni-directional measured RIR. To include the directional characteristics of instruments in the measuring setup, the directivity of the sound source has to be frequency dependent and must be matched to the (measured) directivity of the real instrument. This can be obtained by using a dodecahedron loudspeaker with independently operating systems and an appropriate complex FIR filtering of the frequency response of each driver. The directivity is a result of parameters like magnitude and phase and the interference sum of all systems. To create the appropriate directivity, optimization algorithms are used to achieve minimum error between measured and adapted directivity.

10:15

**4aAA8. Objective measures for evaluating tonal balance of sound fields.** Daiji Takahashi (Dept. of Urban and Environ. Eng., Kyoto Univ., Kyoto Univ. Katsura, Nishikyo-ku, Kyoto 615-8540, Japan, tkhs@archi.kyoto-u.ac.jp), Kanta Togawa (FUJITEC Co., Ltd., Hikone Shiga 522-8588, Japan), and Tetsuo Hotta (YAMAHA Corp., Hamamatsu Shizuoka 430-8650, Japan)

The purpose of this study is to derive objective measures, which can well represent the characteristics of the sound field regarding the tonal balance corresponding to our hearing sense. Two kinds of listening test were conducted in the form of paired comparison, in which subjects were tested using sound fields produced by convoluting some anechoic music sources with some impulse responses. In the first listening test, impulse responses were calculated theoretically for a simple structure of sound field having a direct sound and reflections, and, in the second test, impulse responses were measured at the various seats of existing concert halls. In the latter case, impulse responses which give almost the same reverberance were used for the listening tests. From this investigation, it is found that one objective measure called the DL (deviation of level) has a possibility of an effective measure, which can be used as an appropriate measure for evaluating the tonal balance of sound fields. The index DL is calculated from the data based on the logarithmic scale in both the frequency and magnitude. This fact is not inconsistent with the past findings that human response corresponds to a logarithmic scale of stimulus.

10:35

**4aAA9. Measuring impulse responses containing complete spatial information.** Angelo Farina, Paolo Martignon, Andrea Capra, and Simone Fontana (Industrial Eng. Dept., Univ. of Parma, via delle Scienze 181/A, 43100 Parma, Italy)

Traditional impulse response measurements did capture limited spatial information. Often just omnidirectional sources and microphones are employed. In some cases it was attempted to get more spatial information employing directive transducers: known examples are binaural microphones, figure-of-8 microphones, and directive loudspeakers. However, these approaches are not scientifically based and do not provide an easy way to process and visualize the spatial information. On the other side, psychoacoustics studies demonstrated that “spatial hearing” is one of the dominant factors for the acoustic quality of rooms, particularly for theatres and concert halls. Of consequence, it is necessarily to reformulate the problem entirely, describing the transfer function between a source and a receiver as a time/space filter. This requires us to “sample” the impulse response not only in time, but also in space. This is possible employing spherical harmonics for describing, with a predefined accuracy, the directivity pattern of both source and receiver. It is possible to build arrays of microphones and of loudspeakers, which, by means of digital filters, can provide the required directive patterns. It can be shown how this makes it possible to extract useful information about the acoustical behavior of the room and to make high-quality auralization.

10:55

**4aAA10. Spherical and hemispherical microphone arrays for capture and analysis of sound fields.** Ramani Duraiswami, Zhiyun Li, Dmitry N. Zotkin, and Elena Grassi (Perceptual Interfaces and Reality Lab., Inst. for Adv. Comput. Studies, Univ. of Maryland, College Park, MD 20742)

The capture of the spatial structure of a sound field and analysis is important in many fields including creating virtual environments, source localization and detection, noise suppression, and beamforming. Spherical microphone arrays are a promising development to help achieve such capture and analysis, and have been studied by several groups. We develop a practical spherical microphone array and demonstrate its utility in applications for sound capture, room measurement and for beamforming and tracking. To accommodate equipment failure and manufacturing imprecision we extend their theory to handle arbitrary microphone placement. To handle speech capture and surveillance we describe the development of a new sensor, the hemispherical microphone array. For each array the practical performance follows that predicted by theory. Future applications and improvements are also discussed. [Work supported by NSF.]

11:15

**4aAA11. High-order wave decomposition using a dual-radius spherical microphone array.** Boaz Rafaely, Ilya Balmages, and Limor Eger (Dept. of Elec. and Comput. Eng., Ben-Gurion Univ. of the Negev, Beer-Sheva 84105, Israel)

The acoustic performance of an auditorium is influenced by the manner in which sound propagates from the stage into the seating areas. In particular, the spatial and temporal distribution of early reflections is considered important for sound perception in the auditorium. Previous studies presented measurement and analysis methods based on spherical microphone arrays and plane-wave decomposition that could provide information on the direction and time of arrival of early reflections. This paper presents recent results of room acoustics analysis based on a spherical microphone array, which employs high spherical harmonics order for improved spatial resolution, and a dual-radius spherical measurement array to avoid ill-conditioning at the null frequencies of the spherical Bessel function. Spatial-temporal analysis is performed to produce directional impulse responses, while time-windowed space-frequency analysis is employed to detect direction of arrival of individual reflections. Experimental results of sound-field analysis in a real auditorium will also be presented.

**4aAA12. Impulse response measurement system and its recent applications.** Kazuhiro Takashima, Hiroshi Nakagawa, Natsu Tanaka, and Daiki Sekito (1-21-10, Midori, Sumida-Ku, Tokyo 130-0021, Japan)

Our impulse response measurement system has been developed for ten years. During this decade, the environment related to this measurement has changed significantly. In this article, the features and notes on the measurement system using the sound card, and our brand new system, which is expanded for multichannel inputs, will be presented. Finally, a new technique, which combines multi-channel impulse response measurement and signal processing with microphone array, will be presented. The microphone array was designed for noise analysis for automobile interiors. The array consists of 31 microphones on the surface of an acoustically hard sphere. Moreover, 12 cameras are arranged on the surface of the sphere to take photos. Some applications and future development will be presented.

FRIDAY MORNING, 1 DECEMBER 2006

KOHALA/KONA ROOM, 8:00 TO 11:45 A.M.

## Session 4aAB

### Animal Bioacoustics: Marine Mammal Acoustics I

Paul E. Nachtigall, Chair

*Hawaii Inst. of Marine Biology, P.O. Box 1106, Kailua, HI 96734*

#### Contributed Papers

8:00

**4aAB1. Development of evoked-potential audiometry in odontocetes.**

Alexander Supin (Inst. of Ecology and Evolution, 33 Leninsky prospect, 119071 Moscow, Russia)

Evoked-potential methods are widely used for investigation of hearing in whales, dolphins, and porpoises. For this purpose, mostly the auditory brainstem response (ABR) or rhythmic trains of ABRs, the envelope-following response (EFR), are used. Although very productive, these methods require further elaboration. (i) Traditionally the EFR is provoked by sinusoidally amplitude-modulated tones (SAM). SAM stimuli have narrow frequency band, which makes them little effective to produce the EFR, because response amplitude depends on the stimulus bandwidth. A solution of the problem is the use of trains of short tone pips instead of SAM tones. Such stimuli produce several times higher EFR than SAM tones. This makes the threshold determination much more confident and precise. The effect is achievable at stimulus bandwidths, which still do not influence negatively the precision of attribution of the threshold to a certain frequency. (ii) To extract low-amplitude evoked potentials from noise, the average technique is traditionally used. This operation returns a mean value of averaged traces. Effectively diminishing stationary noise, this method poorly eliminates big artifacts, which may spoil the record even if it appeared once or twice during acquisition. With this respect, computation of the median instead of mean is much more effective.

8:15

**4aAB2. Towards a predictive model of noise-induced temporary threshold shift for an amphibious marine mammal, the California sea lion (*Zalophus californianus*).**

David Kastak, Marla M. Holt, Jason Mulsow, Colleen J. Reichmuth Kastak, Ronald J. Schusterman (UCSC Long Marine Lab., 100 Shaffer Rd., Santa Cruz, CA 95060), and Brandon L. Southall (Natl. Marine Fisheries Service, Silver Spring, MD 20910)

A California sea lion that had previously been tested under water was assessed for noise-induced temporary threshold shift (TTS) in air. One hundred ninety-two controlled exposures of octave-band noise centered at 2.5 kHz were conducted over a 3-year period. The noise was varied in level (to 133 dB SPL *re*: 20  $\mu$ Pa) and duration (to 50 min) to generate a variety of equal sound exposure levels (SELs). Behavioral psychophysics was used to measure hearing sensitivity at 2.5 kHz before, immediately following, and 24 h following noise exposure. The levels of threshold

shifts obtained ranged up to 30 dB. In cases where TTS exceeded 20 dB, thresholds were obtained at regular intervals until recovery occurred. The average slope of the long-term recovery function was 10 dB per log-(minute). Results show that the threshold shifts correlated with SEL; however, the equal-energy trading rule did not apply in all circumstances, with exposure duration contributing more than exposure level. Repeated testing showed no evidence of a permanent threshold shift at 2.5 kHz or octave higher. The amphibious sea lions appear to be equally susceptible to noise in air and under water, provided that the exposure levels are referenced to the subjects thresholds in both media.

8:30

**4aAB3. Electrophysiological investigation of temporal resolution in three pinniped species: Adaptive implications.**

Jason Mulsow and Colleen Reichmuth Kastak (Univ. of California Santa Cruz, Long Marine Lab., 100 Shaffer Rd., Santa Cruz, CA 95060)

Electrophysiological studies of auditory temporal processing in marine mammals have traditionally focused on the role of highly refined temporal resolution in dolphin echolocation. Studies in manatees, however, have found their temporal resolution to be better than expected, leading to speculation that such capabilities are an adaptation for underwater sound localization. This study measured the ability of auditory brainstem responses to follow rhythmic click stimuli in California sea lions (*Zalophus californianus*), harbor seals (*Phoca vitulina*), and northern elephant seals (*Mirounga angustirostris*). Trains of 640-s clicks were presented in air at repetition rates of 125–1500 per second and averaged rate-following responses were recorded. Rate-following responses were detected in both the harbor seal and the sea lion at rates up to 1000 clicks per second, indicating that pinnipeds, like manatees, possess temporal resolution greater than humans but inferior to dolphins. While this finding might support an underwater sound localization hypothesis, comparable results were obtained in preliminary testing of a dog (*Canis familiaris*), suggesting that increased temporal resolution in pinnipeds may not be the result of the evolutionary pressure of an aquatic environment, but rather a result of increased high-frequency hearing essential to mammalian sound localization. [Work supported by NOPP, ONR, and NMFS.]

8:45

**4aAB4. Click and tone-pip auditory evoked potentials in a large marine mammal, the northern elephant seal.** Dorian S. Houser (BIOMIMETICA, 7951 Shantung Dr., Santee, CA 92071) and James J. Finneran (Space and Naval Warfare Systems Ctr., San Diego, CA 92152)

The use of auditory-evoked potentials (AEPs) to study the hearing of mysticete whales is challenged by access to animals, their large size, and proportionately smaller brain relative to odontocetes. One means by which AEP techniques can be adapted to these larger animals is by application to more readily available proxy species. The northern elephant seal (*Mirounga angustirostris*) is a large pinniped, potentially in excess of 2000 kg, with a thick dermis, large skull, relatively small auditory nerve, and a low-frequency vocal communication system. AEP collection in elephant seals provides similar challenges to those of the mysticetes but at a scale that provides a greater opportunity for success. AEP tests were conducted on northern elephant seals at Año Nuevo State Reserve, the natural haul-out site of the elephant seal. Subjects were chemically immobilized with tiletamine/zolazepam and chemical restraint was maintained with bolus injections of ketamine. Click-evoked potentials were collected from four weanling and two adult male elephant seals and tone-pip-evoked potentials were collected from a 2-year-old female. Results demonstrate that AEPs can be recorded from large pinniped species, providing a step towards the application of similar techniques to larger cetacean species.

9:00

**4aAB5. Acoustic field measurements and bottlenose dolphin hearing thresholds using single-frequency and frequency-modulated tones.**

James J. Finneran (U.S. Navy Marine Mammal Program, SPAWAR/SYSCEN San Diego, Code 2351, 49620 Beluga Rd., San Diego, CA 92152, james.finneran@navy.mil) and Carolyn E. Schlundt (EDO Professional Services, San Diego, CA 92110)

Studies of underwater hearing are often hampered by the behavior of sound waves in small experimental tanks. At lower frequencies, tank dimensions are often not sufficient for free-field conditions, resulting in large spatial variations of sound pressure. These effects may be mitigated somewhat by increasing the frequency bandwidth of the sound stimulus, so effects of multipath interference average out over many frequencies. In this study, acoustic fields and bottlenose dolphin (*Tursiops truncatus*) hearing thresholds were compared for pure-tone and frequency-modulated stimuli. Experiments were conducted in a vinyl-walled, seawater-filled pool approximately 4×5×1.5 m. Sound stimuli consisted of 500-ms tones at 13 carrier frequencies between 1 and 100 kHz. Frequency-modulated stimuli featured both linear and sinusoidal modulating waveforms with 5%, 10%, and 20% bandwidths. Acoustic fields were measured (without the dolphin present) at three depths over a 60×65-cm grid with a 5-cm spacing. Hearing thresholds were measured using a behavioral response paradigm and up/down staircase technique. Frequency-modulated stimuli with a 10% bandwidth resulted in significant improvements to the sound field without substantially affecting the dolphins hearing thresholds. [Work supported by ONR.]

9:15

**4aAB6. Hearing frequency selectivity in four species of toothed whales as revealed by the evoked-potential method.**

Vladimir Popov (Inst. of Ecology and Evolution, 33 Leninsky Prosp., 119071 Moscow, Russia popov\_vl@sevin.ru)

Frequency tuning curves were obtained using a tone-tone simultaneous masking paradigm in conjunction with the evoked potential recording. The masker was a continuous tone and the test was a sinusoidal amplitude-modulated (SAM) tonal signal, which evoked the envelope following response (EFR). The EFR was recorded in unanaesthetized animals from a head surface with the use of suction-cup electrodes. The obtained tuning curves featured very sharp tuning with Q(ERB) (quality estimated by the equivalent rectangular bandwidth) from 35 in *Tursiops truncatus* to nearly 50 in *Delphinapterus leucas*. This acuteness is several times better than in humans and many animals. The Q(ERB) dependence on probe frequency could be approximated by regression lines with a slope from 0.18 in *Tur-*

*siops truncatus* to 0.83–0.86 in *Phocoena phocoena* and *Neophocoena phocoenoides*. Thus, the frequency representation in the odontocete auditory system may be either near constant quality (in *Tursiops*) or near constant bandwidth (in porpoises). [Work supported by The Russian Foundation for Basic Research and Russian President Grant.]

9:30

**4aAB7. Growth and recovery of temporary threshold shifts in a dolphin exposed to midfrequency tones with durations up to 128 s.**

Carolyn E. Schlundt (EDO Professional Services, 3276 Rosecrans St., San Diego, CA 92110, carolyn.melka@edocorp.com), Randall L. Dear (Sci. Applications Intl. Corp., San Diego, CA 92110), Donald A. Carder, and James J. Finneran (Space and Naval Warfare Systems Ctr., San Diego, San Diego, CA 92152)

Auditory thresholds at 4.5 kHz were measured in a bottlenose dolphin (*Tursiops truncatus*) before and after exposure to midfrequency tones at 3 kHz. Experiments were conducted in relatively quiet pools with low ambient noise levels at frequencies above 1 kHz. Behavioral hearing tests allowed for thresholds to be routinely measured within 4 min postexposure, and tracked recovery for at least 30 min postexposure. Exposure durations ranged from 4 to 128 s at sound pressure levels ranging from 149 to 200 dB re: 1 μPa. Sound exposure levels ranged from 155 to 217 dB re: 1 μPa<sup>2</sup>/s. Temporary threshold shifts at 4 min postexposure (TTS<sub>4</sub>) of up to 23 dB were observed. All thresholds recovered to baseline and pre-exposure levels, most within 30 min of exposure. [Work supported by the U.S. ONR.]

9:45

**4aAB8. Auditory brainstem response recovery rates during double-pulse presentation in the false killer whale (*Pseudorca crassidens*): A mechanism of automatic gain control?**

Paul E. Nachtigall (Marine Mammal Res. Program, Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734), Alexander Ya. Supin (Russian Acad. of Sci., Moscow, Russia), and Marlee Breese (Hawaii Inst. of Marine Biol., Kailua, HI 96734)

The outgoing echolocation pulse and the return echo response can be approximately examined in the auditory system of an echolocating animal by presenting two pulses and determining the forward-masking effect of the first pulse on the response to the second pulse using auditory-evoked potential procedures. False killer whale, *Pseudorca crassidens*, auditory brainstem responses (ABR) were recorded using a double-click stimulation paradigm specifically measuring the recovery of the second (test) response (to the second click) as a function of the length of the interclick interval (ICI) following various levels of the first (conditioning) click. At all click intensities, the slopes of the recovery functions were almost constant: 0.60.8 V per ICI decade. Therefore, even when the conditioning-to-click level ratio was kept constant, the duration of recovery was intensity dependent: the higher intensity the longer the recovery. The conditioning-to-test-click level ratio strongly influenced the recovery time: the higher the ratio, the longer the recovery. This dependence was nearly linear, using a logarithmic ICI scale with a rate of 2530 dB per ICI decade. These data were used for modeling the interaction between the emitted click and the echo in the auditory system during echolocation.

10:00–10:15 Break

10:15

**4aAB9. Temporary threshold shifts in the bottlenose dolphin (*Tursiops truncatus*), varying noise duration and intensity.**

T. Aran Mooney (Dept. of Zoology and Hawaii Inst. of Marine Biol., Univ. of Hawaii, 46-007 Lilipuna Rd., Kaneohe, HI 96744), Paul E. Nachtigall, Whitlow W. L. Au, Marlee Breese, and Stephanie Vlachos (Univ. of Hawaii, Kaneohe, HI 96744)

There is much concern regarding increasing noise levels in the ocean and how it may affect marine mammals. However, there is a little information regarding how sound affects marine mammals and no published

4a FRI. AM

data examining the relationship between broadband noise intensity and exposure duration. This study explored the effects of octave-band noise on the hearing of a bottlenose dolphin by inducing temporary hearing threshold shifts (TTS). Sound pressure level (SPL) and exposure duration were varied to measure the effects of noise duration and intensity. Hearing thresholds were measured using auditory evoked potentials before and after sound exposure to track and map TTS and recovery. Shifts were frequency dependent and recovery time depended on shift and frequency, but full recovery was relatively rapid, usually within 20 min and always within 40 min. As exposure time was halved, TTS generally occurred with an increase in noise SPL. However, with shorter, louder noise, threshold shifts were not linear but rather shorter sounds required greater sound exposure levels to induce TTS, a contrast to some published literature. From the data a novel algorithm was written that predicts the physiological effects of anthropogenic noise if the intensity and duration of exposure are known.

10:30

**4aAB10. Estimates of bio-sonar characteristics of a free-ranging Ganges river dolphin.** Tamaki Ura, Harumi Sugimatsu, Tomoki Inoue (Underwater Technol. Res. Ctr., Inst. of Industrial Sci., Univ. of Tokyo, 4-6-1 Komaba, Meguro, Tokyo 153-8505, Japan), Rajendar Bahl (IIT Delhi, New Delhi 110016, India), Junichi Kojima (KDDI R&D Labs. Inc., Saitama 356-8502, Japan), Tomonari Akamatsu (Fisheries Res. Agency, Ibaraki 314-0408, Japan), Sandeep Behera (Freshwater & Wetlands Prog., New Delhi 110003, India), Ajit Pattnaik, Muntaz Kahn (Chilika Dev. Authority, Orissa, India), Sudhakar Kar, Chandra Sekhar Kar (Off. of the Princip. CCF (Wildlife) & Chief Wildlife Warden, Bhubaneswar 751012, India), Tetsuo Fukuchi, Hideyuki Takashi (System Giken Co. Ltd., Kanagawa 253-0085, Japan), and Debabrata Swain (Simpilpal Biosphere and Tiger Reserve, Orissa, India)

This paper reports the first known studies of the bio-sonar characteristics of an isolated free-ranging Ganges river dolphin, *Platanista gangetica*. The animal preferred to roam in a deeper tract of the otherwise shallow river. The click sounds of the dolphin were recorded over a period of 2 days on a 3.2-meter-long high-frequency hydrophone array composed of three hydrophones forming an equispaced linear array and another two hydrophones in conjunction with the central hydrophone forming an SSBL triangular array in a plane perpendicular to the array axis. The array was deployed both in horizontal and vertical configurations. The array structure provided 3-D measurements of the source location through measurement of the interelement time delay. Bio-sonar characteristics such as click duration, bandwidth, and interclick intervals in click trains have been reported. Measurements of dolphin track and the relative click levels on the array hydrophones have been used to obtain a preliminary characterization of the animal's beam pattern.

10:45

**4aAB11. Discriminating between the clicks produced by a bottlenose dolphin when searching for and identifying an object during a search task.** Sandra Bohn, Stan Kuczaj (Univ. of Southern Mississippi, 118 College Dr., #5025, Hattiesburg, MS 39406, sandra.bohn@usm.edu), and Dorian Houser (BIOMIMETICA, Santee, CA 92071)

Clicks collected from an echolocating bottlenose dolphin completing a search task were compared in order to determine if the clicks produced when the dolphin was acquiring the target differed from the clicks produced when the dolphin was searching for the target. The clicks produced by a free-swimming dolphin completing the search task were recorded using a biosonar measurement tool (BMT), an instrumentation package carried by the dolphin that collected both the outgoing clicks and the returning echoes. A discriminant function analysis classified the clicks as search or acquisition using the variables of peak-to-peak amplitude, duration, peak frequency, center frequency, and bandwidth. The acquisition clicks were classified more accurately than the search clicks. Acquisition clicks and search clicks were significantly different across all five of the variables. These results suggest that the clicks produced by bottlenose dolphins acquiring a target are different than those produced by dolphins searching for a target.

11:00

**4aAB12. Echo highlight amplitude and temporal difference resolutions of an echolocating *Tursiops truncatus*.** Mark W. Muller, Whitlow W. L. Au, Paul E. Nachtigall, Marlee Breese (Marine Mammal Res. Program, Hawai'i Inst. of Marine Biol., 46-007 Lilipuna Rd., Kaneohe, HI 96744), and John S. Allen III (Univ. of Hawai'i at Manoa, Honolulu, HI 96822)

A dolphin's ability to discriminate targets may greatly depend on the relative amplitudes and the time separations of echo highlights within the received signal. Previous experiments with dolphins have varied the physical parameters of targets, but did not fully investigate how changes in these parameters corresponded with the composition of the scattered acoustic waveforms and the dolphin's subsequent response. A novel experiment utilizes a phantom echo system to test a dolphin's detection response of relative amplitude differences of secondary echo highlights and the time separation differences of all the echo highlights both within and outside the animal's integration window. By electronically manipulating these echoes, the underlying acoustic classification cues can be more efficiently investigated. In the first study, the animal successfully discriminated between a standard echo signal and one with the middle highlight amplitude at  $-7$  dB. When the middle highlight amplitude was raised to  $-6$  dB, the animal's discrimination performance radically dropped to 65%. This study suggests the animal may not be as sensitive to the secondary echo highlights as previously proposed. The experiments were repeated for the trailing highlight amplitude and the time separations between the primary and middle highlights and the middle and trailing highlights.

11:15

**4aAB13. A background noise reduction technique for improving false killer whale (*Pseudorca crassidens*) localization.** Craig R. McPherson, Owen P. Kenny, Phil Turner (Dept. of Elec. and Comput. Eng., James Cook Univ., Douglas 4811, Queensland, Australia), and Geoff R. McPherson (Queensland Dept. of Primary Industries and Fisheries, Cairns, 4870, Queensland Australia)

The passive localization of false killer whales (*Pseudorca crassidens*) in acoustic environments comprised of discontinuous ambient, anthropogenic, and animal sounds is a challenging problem. A background noise reduction technique is required to improve the quality of sampled recordings, which will assist localization using auditory modeling and signal correlation at extended ranges. The algorithm developed meets this requirement using a combination of adaptive percentile estimation, a median-based tracker, and Gaussian windowing. The results indicate successful improvement of the signal-to-noise ratio, and consequently a significant increase in the detection range of false killer whales in acoustically complex environments.

11:30

**4aAB14. Analysis of Australian humpback whale song using information theory.** Jennifer L. Miksis-Olds, John R. Buck (School for Marine Sci. and Technol., Univ. of Massachusetts Dartmouth, New Bedford, MA 02744, jmiksis@umassd.edu), Michael J. Noad (Univ. of Queensland, St. Lucia, QLD 4072 Australia), Douglas H. Cato (Defence Sci. & Tech. Org., Pyrmont, NSW 2009 Australia), and Dale Stokes (Scripps Inst. of Oceanogr., La Jolla, CA 92093)

Songs produced by migrating whales were recorded off the coast of Queensland, Australia over 6 consecutive weeks in 2003. Approximately 50 songs were analyzed using information theory techniques. The average length of the songs estimated by correlation analysis was approximately 100 units, with song sessions lasting from 300 to over 3100 units. Song entropy, a measure of structural constraints and complexity, was estimated using three different methodologies: (1) the independently identically distributed model; (2) first-order Markov model; and (3) the nonparametric sliding window match length (SWML) method, as described in Suzuki *et al.* [J. Acoust. Soc. Am. **119**, 1849 (2006)]. The analysis finds the songs of migrating Australian whales are consistent with the hierarchical structure proposed by Payne and McVay (Science **173**, 585–597 (1971)), and

recently confirmed by Suzuki *et al.* for singers on the breeding grounds. Both the SWML entropy estimates and the song lengths for the Australian singers were lower than that reported by Suzuki *et al.* for Hawaiian whales in 1976–1978. These lower SWML entropy values indicate a

higher level of predictability within songs. The average total information in the Australian sequence of song units was approximately 35 bits/song. Aberrant songs (10%) yielded entropies similar to the typical songs. [Sponsored by ONR and DSTO.]

FRIDAY MORNING, 1 DECEMBER 2006

KAHUKU ROOM, 7:55 A.M. TO 12:00 NOON

### Session 4aBB

## Biomedical Ultrasound/Bioresponse to Vibration: Interaction of Cavitation Bubbles with Cells and Tissue

John S. Allen, Cochair

*Univ. of Hawaii, Dept. of Mechanical Engineering, 2540 Dole St., Honolulu, HI 96822*

Yoshiki Yamakoshi, Cochair

*Gunma Univ., Faculty of Engineering, 1-5-1 Tenjin-cho, Kiryu-shi, Gunma 376-8515, Japan*

Chair's Introduction—7:55

### Invited Papers

8:00

**4aBB1. Ultra-high-speed imaging of bubbles interacting with cells and tissue.** Michel Versluis, Philippe Marmottant, Sascha Hilgenfeldt, Claus-Dieter Ohl (Phys. of Fluids, Univ. of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands), Chien T. Chin, Annemieke van Wamel, Nico de Jong (Erasmus MC, 3000 DR Rotterdam, The Netherlands), and Detlef Lohse (Univ. of Twente, 7500 AE Enschede, The Netherlands)

Ultrasound contrast microbubbles are exploited in molecular imaging, where bubbles are directed to target cells and where their high-scattering cross section to ultrasound allows for the detection of pathologies at a molecular level. In therapeutic applications vibrating bubbles close to cells may alter the permeability of cell membranes, and these systems are therefore highly interesting for drug and gene delivery applications using ultrasound. In a more extreme regime bubbles are driven through shock waves to sonoporate or kill cells through intense stresses or jets following inertial bubble collapse. Here, we elucidate some of the underlying mechanisms using the 25-Mfps camera Brandaris128, resolving the bubble dynamics and its interactions with cells. We quantify acoustic microstreaming around oscillating bubbles close to rigid walls and evaluate the shear stresses on nonadherent cells. In a study on the fluid dynamical interaction of cavitation bubbles with adherent cells, we find that the nonspherical collapse of bubbles is responsible for cell detachment. We also visualized the dynamics of vibrating microbubbles in contact with endothelial cells followed by fluorescent imaging of the transport of propidium iodide, used as a membrane integrity probe, into these cells showing a direct correlation between cell deformation and cell membrane permeability.

8:20

**4aBB2. Sonoporation: Mechanisms of cell membrane perforation and rapid resealing.** Nobuki Kudo and Katsuyuki Yamamoto (Grad. School of Information Sci. and Technol., Hokkaido Univ., Sapporo 060-0814 Japan, kudo@bme.ist.hokudai.ac.jp)

Sonoporation is a technique for making membrane perforation by exposure of cells to ultrasound, and it is an attractive method for realizing nonvirus gene transfection. A continuous or quasicontinuous wave is frequently used for this technique because a higher duty ratio gives higher efficiency of sonoporation. Addition of microbubbles during insonification greatly improves the efficiency of sonoporation, and, especially when microbubbles exist in the vicinity of the cells, ultrasound pulses from diagnostic ultrasound equipment can cause sonoporation. In this study, we examined sonoporation induced by single-shot pulsed ultrasound and the role of microbubbles in induction of cell membrane perforation. Bubble behavior and cell membrane damage were observed using a high-speed camera and light and scanning electron microscopes. Results of observation showed that mechanical stress induced by bubble motion could cause cell membrane perforation. We also studied repair of the perforation using a fluorescence microscope and found that the membrane of mammalian cells has the ability to reseal the perforation within several seconds. [Research partially supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture, Japan.]

8:40

**4aBB3. Quantitative imaging of tumor blood flow with contrast ultrasound.** Peter N. Burns, Raffi Karshafian, and John Hudson (Dept. Medical Biophys., 2075 Bayview Ave., Toronto ON, M4N 3M5, Canada)

The point at which a solid cancer develops its own blood supply marks the onset of malignant progression. This process, known as angiogenesis, makes oxygen and nutrients available for growth and provides a path for metastatic spread. Angiogenesis is not only of interest as a diagnostic hallmark of malignancy, but also as a target for new therapeutic strategies. Assessing antiangiogenic therapies noninvasively poses problems—flow velocities ( $<1$  mm/s) and vessel diameters ( $<50$   $\mu\text{m}$ ) are below resolutions of direct imaging. Vessels are disorganized without the tree-like structure of normal vasculature. We have investigated the potential role of

microbubble disruption-replenishment flow measurement in monitoring antivasular treatment of an animal tumor. The currently used monexponential model incorrectly considers the vasculature a perfect mixing chamber. Simple fractal models of the circulation provide a distribution of vessel diameters which, combined with the geometry of the disruption and detection beams, produce better models of replenishment following acoustic bubble disruption. These not only measure flow, but also predicts differences between organized and disorganized circulations, even with equal flow and vascular volume. If detectable, such differences might be used to characterize vascular organization below the resolution limit of an ultrasound image.

9:00

**4aBB4. Dynamics of laser-trapped microbubbles.** Hiroyuki Takahira (Dept. of Mech. Eng., Osaka Prefecture Univ., 1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan)

A laser-trapping method is utilized for microbubbles. Bubbles of the order of 10 microns in diameter are trapped and manipulated successfully using a dry objective lens with large working distance. The growth or shrinkage of a laser-trapped microbubble and the merger of microbubbles are observed with a high-speed camera to investigate the influence of gas diffusion on the stability of microbubbles. Two kinds of equilibrium radii are found for shrinking microbubbles. The first one is related to the equilibrium surface concentration of surfactant. The other is related to the decrease of the surface tension due to the compression of the surface area at the maximum surfactant concentration. The simulations in which the dynamic surface tension is considered are in good agreement with the experiments. The laser trapping technique is also applied to the motion of a microbubble in a shear flow. It is shown that the bubble escapes from the laser trap being repelled by the optical force in the shear flow. There is overall agreement between the experiments and the simulations in which the buoyancy force, the fluid dynamic forces, and the optical force are taken into account.

9:20

**4aBB5. Novel methods of micro-object trapping by acoustic radiation force.** Yoshiki Yamakoshi (1-5-1 Tenjin-cho, Kiryushi, Gunma 376-8515 Japan, yamakosi@el.gunma-u.ac.jp)

It is expected that micro object trapping by acoustic radiation force is a useful method in future drug delivery systems in order to concentrate the payloads at desired position. In this paper, two novel methods of micro object trapping are presented. First method is micro object trapping by seed bubbles. This method uses seed bubbles, which have higher sensitivity to the ultrasonic wave, in order to trap micro objects, which are difficult to trap by conventional methods due to low volumetric oscillation under the ultrasonic wave. The Bjerkne's force, which is produced by a secondary wave radiated from the seed bubbles, traps the target objects making bi-layer seed bubbletarget object mass. The Second method is micro bubble trapping by bubble nonlinear oscillation. Two ultrasonic waves with different frequencies (pumping and control waves) are introduced simultaneously. The frequency of the control wave is set to a harmonic frequency of the pumping wave. If the bubbles flow into the cross area of two waves, nonlinear oscillation by high intensity pumping wave generates the Bjerkne's force, producing multiple traps with narrow separation along the control wave propagation direction. In order to demonstrate these methods, experiments using an ultrasonic wave contrast agent are shown.

9:40

**4aBB6. Mechanical properties of HeLa cells at different stages of cell cycle by time-resolved acoustic microscope.** Pavel V. Zinin (School of Ocean and Earth Sci. and Technol., Univ. of Hawaii, 2525 Correa Rd., Honolulu, HI 96822-2219), Eike C. Weiss, Pavlos Anastasiadis, and Robert M. Lemor (Fraunhofer Inst. for Biomed. Technol., St. Ingbert, Germany)

Scanning acoustic microscopy (SAM), particularly time-resolved acoustic microscopy, is one of the few techniques for study of the mechanical properties of only the cell's interior, cytosol and nucleus. Unfortunately, time-resolved acoustic microscopes typically do not provide sufficient resolution to study the elasticity of single cells. We demonstrate that the high-frequency, time-resolved acoustic microscope developed at the Fraunhofer Institute for Biomedical Technology (IBMT), Germany, is capable of imaging and characterizing elastic properties of micron size structures in cell's cytoskeleton with a theoretical resolution limit of 10 m/s for sound speed measurements. Measurements were performed on cells of the HeLa cell line derived from human cervix carcinoma. SAM measurements of the sound speed of adherent HeLa cells at different states of the cell cycle were conducted. They yielded an average value of 1540 m/s. B-Scan images of HeLa cells at different states of the cell cycle show distinct patterns inside the cell. A method for estimating sound attenuation inside HeLa cells is outlined as such a method is critical for the determination of a cell's viscoelasticity. [Work supported by Alexander von Humboldt Foundation and the European Framework Program 6, Project "CellProm."]

10:00–10:10 Break

10:10

**4aBB7. Assessment of shock wave lithotripters via cavitation potential.** Jonathan I. Ilroeta, Andrew J. Szeri (UC Berkeley, 6119 Etcheverry Hall, M.S. 1740, Berkeley, CA 94720-1740), Yufeng Zhou, Georgii Sankin, and Pei Zhong (Duke Univ., Durham, NC 27708-0300)

An analysis of bubbles in elastic media has been made in order to characterize shock wave lithotripters by gauging the potential for cavitation associated with the lithotripter shock wave (LSW). The method uses the maximum radius achieved by a bubble subjected to a LSW as the key parameter that defines the potential damage a lithotripter could cause at any point in the domain. The maximum radius is determined by an energy analysis. A new index—similar in spirit to the Mechanical Index of Holland and Apfel for diagnostic ultrasound—is proposed for use in gauging the likelihood of cavitation damage.

10:30

**4aBB8. Formation of water pore in a bilayer induced by shock wave: Molecular dynamics simulation.** Kenichiro Koshiyama, Takeru Yano, Shigeo Fujikawa (Lab. of Adv. Fluid Mech., Hokkaido Univ., Sapporo 060-8628, Japan, koshi@ring-me.eng.hokudai.ac.jp), and Tetsuya Kodama (Tohoku Univ., Aobaku, Sendai 980-8575, Japan)

The irradiation of a shock wave or ultrasound with micro-bubbles has the potential to make transient pores on cell membranes. Although such pores are believed to contribute to the molecular delivery thorough the membrane, the detailed mechanisms of the pore formation with shock waves and the subsequent molecular delivery through the pores into cells are still unclear. To investigate the mechanism at a molecular level, the molecular dynamics simulations of the interaction of the shock wave with a lipid bilayer are conducted. The water penetration into the hydrophobic region by the shock wave is observed in picoseconds. As a next step, structural changes of the bilayer containing water molecules in the hydrophobic region are investigated. The water pore is formed in 3 ns when the large number of water molecules is inserted. The lifetime of the water pore is more than 70 ns. The radius of the water pore is ca. 1.0 nm, which is three times larger than the Stoke's radius of a typical anticancer drug (5FU). Finally, the diffusion of the anticancer drug in the water pore is investigated.

10:45

**4aBB9. Ultrasonic spore lysis and the release of intracellular content in a microfluidic channel.** Oana C. Marina, Michael D. Ward, John M. Dunbar, and Gregory Kaduchak (MPA-11, Los Alamos Natl. Lab., P.O. Box 1663, MS D-429, Los Alamos, NM, 87545)

Ultrasonic lysis of suspended spores in a microfluidic channel is a promising alternative to conventional spore disruption techniques that include bead beating as the spore lysis gold standard. Our overall research goal is to obtain an automated detection system with complete sample preparation and lysis steps in a microfluidic channel. Previously, much work in this area has focused on organism viability rather than the release of intracellular material. Our research focuses on quantifying the amount of intracellular content (e.g., DNA, proteins, etc.) that is released by acoustic lysis for detection by the sensor. Elucidating the efficacy of acoustics on the release of intracellular material requires reliable methods to quantify the released intracellular content (nucleic acids and proteins). The device used for lysing spores consists of a microfluidic chamber with one acoustically active wall. The chamber depths are in the range of 100–200  $\mu$ m. Channels tested in the 70-kHz to 1-MHz frequency range show that the efficiency of intracellular release depends on the operating frequency of the device and the properties (concentration, composition) of the spore suspensions. Experimental results on viability and released intracellular content are discussed. [Work supported by LANL LDRD.]

11:00

**4aBB10. The correlation between cavitation noise power and bubble-induced heating in high-intensity focused ultrasound.** Caleb H. Farny, Tianming Wu, R. Glynn Holt, and Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, cfarny@bu.edu)

It has been established that inertial cavitation is responsible for elevated heating during high-intensity focused ultrasound (HIFU) application for certain intensity regimes. The contribution of bubble-induced heating can be an important factor to consider, as it can be several times that expected from absorption of the primary ultrasound energy. Working in agar-graphite tissue phantoms with a 1.1-MHz HIFU transducer, an embedded type-E thermocouple, and a 15-MHz passive cavitation detector (PCD), the temperature and cavitation signal near the focus were measured for 5-s continuous wave HIFU insonations. The measured temperature was corrected for heating predicted from the primary ultrasound absorption and the transient thermocouple viscous heating artifact to isolate the temperature rise from the bubble activity. We have found that the

temperature rise induced from the bubble activity correlates well with the instantaneous cavitation noise power as indicated by the mean square voltage output of the PCD. The results suggest that careful processing of the cavitation signals could serve as a proxy for measuring the heating contribution from inertial cavitation. [Work supported by the Dept. of the Army (Award No. DAMD17-02-2-0014) and the Center for Subsurface Sensing and Imaging Systems (NSF ERC Award No. EEC-9986821).]

11:15

**4aBB11. Membrane permeabilization of adherent cells with laser-induced cavitation bubbles.** Rory Dijkink, Claus-Dieter Ohl (Phys. of Fluids, Univ. of Twente, Postbus 217, 7500 AE Enschede, The Netherlands), Erwin Nijhuis, Sèverine Le Gac (Univ. of Twente, 7500 AE Enschede, The Netherlands), and István Vermes (Medical Spectrum Twente Hospital Group, 7500 KA Enschede, The Netherlands)

Strongly oscillating bubbles close to cells can cause the opening of the cell's membrane, thus to stimulate the uptake of molecules from the exterior. However, the volume oscillations of bubbles induce complex fluid flows, especially when bubble-bubble interaction takes place. Here, we report on an experiment where a single cavitation bubble is created close to a layer of adherent HeLa cells. The interaction distance between the bubble and the cell layer is controlled by adjusting the focus of the pulsed laser light, which creates the cavitation bubble. The dynamics of the bubble and the cells is recorded with high-speed photography. The poration of the cells is probed with different fluorescent stains to distinguish viable and permanent poration and programmed cell death (apoptosis). Quantitative data are presented as a function of the radial distance from the stagnation point. Our main finding is the importance of the asymmetrical collapse and the high-speed jet flow: After impact of the jet onto the substrate a strong boundary layer flow is responsible for shearing the cells.

11:30

**4aBB12. Antitumor effectiveness of cisplatin with ultrasound and nanobubbles.** Tetsuya Kodama, Yukiko Watanabe, Kiyoe Konno, Sachiko Horie (Res. Organization, Tohoku Univ., 2-1 Seiryomachi, Aoba-ku, Sendai, Miyagi 980-8575, Japan), Atsuko Aoi (Tohoku Univ., Sendai 980-8575, Japan), Geroges Vassaux (Bart's and The London School of Medicine and Dentistry, UK), and Shiro Mori (Tohoku Univ. Hospital, Sendai 980-8575, Japan)

The potentiation of antitumor effect of cis-diamminedichloroplatinum (II), cisplatin, with ultrasound (1 MHz, 0.6 MPa) and lipid-shelled nanobubbles *in vitro* (EMT6, C26, MCF7, A549) and *in vivo* on s.c. tumor in mice (HT29-expressing luciferase) were evaluated. *In vitro* and *in vivo* antitumor effects were measured by an MTT assay and a real-time *in vivo* imaging, respectively. The effective antitumor effect was seen both *in vitro* and *in vivo* when ultrasound and nanobubbles were used, while other treatment groups with cisplatin with ultrasound did not show the effectiveness. The antitumor effect was not attributed to necrosis but apoptosis, which was confirmed by increase in the activity of the pro-apoptosis signal caspase-3 and Bax. In conclusion, the combination of ultrasound and nanobubbles with cisplatin is an effective chemotherapy of solid tumors and may prove useful in clinical application.

11:45

**4aBB13. Sonoporation by single-shot pulsed ultrasound with microbubbles—Little effect of sonochemical reaction of inertial cavitation.** Kengo Okada, Nobuki Kudo, and Katsuyuki Yamamoto (Grad. School of Information Sci. and Technol., Hokkaido Univ., Kita 14 Nishi 9, Kita-ku, Sapporo 060-0814, Japan)

Sonoporation is a technique for introducing large molecules into a cell by exposure to ultrasound, and it has a potential application for gene transfection. Although continuous-wave ultrasound is generally used for this technique, we have been using single-shot pulsed ultrasound with microbubbles. To determine the contribution of the sonochemical effect of

inertial cavitation under the condition of single-shot exposure, we compared rates of cell membrane damage in the presence and absence of a free radical scavenger (cysteamine, 5 mM). Cells with microbubbles in their vicinity were exposed to pulsed ultrasound of 1.1 MPa in negative peak pressure under microscopic observation, and the numbers of total and damaged cells in the view field were counted. The damage rates were  $8.1 \pm 4.0\%$  and  $10.3 \pm 6.3\%$  in the presence ( $n=17$ ) and absence ( $n$

$=25$ ) of the scavenger, respectively, and the average number of total cells was  $772 \pm 285$ . Since there was no significant difference, we concluded that the cell membrane damage observed in our exposure condition was not caused by the sonochemical effect but by the mechanical effect of inertial cavitation. [Research was supported by a grant-in-aid for scientific research from the Ministry of Education, Science, Sports and Culture, Japan.]

FRIDAY MORNING, 1 DECEMBER 2006

OAHU ROOM, 8:00 TO 11:50 A.M.

## Session 4aEA

### Engineering Acoustics and ASA Committee on Standards: Developments in Microphones: Calibrations, Standards, and Measures

George S. K. Wong, Cochair

*National Research Council, Inst. for National Measurement Standards, 1500 Montreal Rd., Ottawa, Ontario K1A 0R6, Canada*

Masakazu Iwaki, Cochair

*NHK Science and Technology Research Labs., 1-10-11 Kinuta, Setagaya-ku, Tokyo 157-8510, Japan*

Chair's Introduction—8:00

#### Invited Papers

8:05

**4aEA1. Current developments at the National Institute for Standards and Technology in pressure calibration of laboratory standard microphones.** Victor Nedzelnitsky, Randall P. Wagner, and Steven E. Fick (National Inst. of Standards and Technol. [NIST], 100 Bureau Dr., Stop 8220, Gaithersburg, MD 20899-8220, victor.nedzelnitsky@nist.gov)

Current research effort aims at improving the apparatus and methods for determining the pressure sensitivities of IEC types LS1Pn and LS2aP laboratory standard microphones. Among the improvements that are being systematically incorporated in an evolving test bed is the capability to operate at adjustable power line frequencies other than the usual 60 Hz. Suitable choices of line frequency relative to frequencies of calibration and adjustable bandpass filter characteristics can be used to improve the signal-to-noise ratios of measurements performed near the usual line frequency and its first few harmonics. This can enable the use of relatively large volume couplers for which uncertainties in microphone front cavity volume and equivalent volume, capillary tube effects, and heat conduction corrections have a lesser influence than they have for small-volume couplers. Another improvement aims to control and to stabilize the ambient static pressure during microphone calibrations, to reduce or eliminate the effects of barometric pressure fluctuations on these calibrations.

8:25

**4aEA2. Free-field reciprocity calibration of laboratory standard (LS) microphones using a time selective technique.** Knud Rasmussen and Salvador Barrera-Figueroa (Danish Primary Lab. of Acoust. (DPLA), Danish Fundamental Metrology, Danish Tech. Univ., Bldg. 307, 2800 Kgs., Lyngby, Denmark)

Although the basic principle of reciprocity calibration of microphones in a free field is simple, the practical problems are complicated due to the low signal-to-noise ratio and the influence of cross talk and reflections from the surroundings. The influence of uncorrelated noise can be reduced by conventional narrow-band filtering and time averaging, while correlated signals like cross talk and reflections can be eliminated by using time-selective postprocessing techniques. The technique used at DPLA overcomes both these problems using a B&K Pulse analyzer in the SSR mode (steady state response) and an FFT-based time-selective technique. The complex electrical transfer impedance is measured in linear frequency steps from a few kHz to about three times the resonance frequency of the microphones. The missing values at low frequencies are estimated from a detailed knowledge of the pressure sensitivities. Next an inverse FFT is applied and a time window around the main signal is used to eliminate cross talk and reflections. Finally, the signal is transformed back to the frequency domain and the free field sensitivities calculated. The standard procedure at DPLA involves measurements at four distances and the repeatability of the calibrations over time is within  $\pm 0.03$  dB up to about 1.5 times the resonance frequency of the microphones.

8:45

**4aEA3. Microphone calibration by comparison.** George S. K. Wong (Acoust. Standards, Inst. for Natl. Measurement Standards, National Res. Council Canada, Ottawa, ON K1A 0R6, Canada)

The absolute method of microphone calibration by the reciprocity method (IEC 61094-2 1992) provides the highest accuracy of approximately 0.04 to 0.05 dB, and the procedure requires three changes of microphone in the “driver-receiver combination” that needs approximately 1 to 2 days. The system capital cost is relatively high. The NRC interchange microphone method for microphone calibration by comparison has been adopted internationally by the International Electrotechnical Commission as IEC 61094-5 (2001-10). With this method, the test microphone is compared with a reference microphone calibrated by the reciprocity method and the procedure requires approximately 3 h. The uncertainty of the comparison method is between 0.08 to 0.09 dB, which satisfies most industrial needs.

9:05

**4aEA4. Development of a laser-pistonphone for an infrasonic measurement standard.** Ryuzo Horiuchi, Takeshi Fujimori, and Sojun Sato (Natl. Metrology Inst. of Japan (NMIJ), AIST, Tsukuba Central 3, 1-1-1 Umezono, Tsukuba, 305-8563, Japan)

Acoustical standards for audio frequencies are based on pressure sensitivities of laboratory standard microphones calibrated using a coupler reciprocity technique. There is a growing need to extend the frequency range downward for reliable infrasonic measurement. The reciprocity technique, however, has limitations on low-frequency calibration (1–20 Hz) because signal-to-noise ratio deteriorates and a sound leak occurs from capillary tubes that equalize the static pressure inside and outside of the coupler. These factors rapidly increase the measurement uncertainty as the frequency is lowered. NMIJ has therefore recently developed a laser-pistonphone prototype, which enables precise calibration of microphones at low frequencies. Compared with the reciprocity technique, the laser-pistonphone produces a higher sound pressure within a cavity by the sinusoidal motion of a piston and has a significantly improved signal-to-noise ratio. Sound pressure is calculated from the piston displacement, which is determined via a Michelson interferometer. A test microphone is inserted into the cavity, exposed to the sound pressure, and its open-circuit voltage is measured. Static pressure equalization is realized through the gap between the piston and its guide. Careful design of the dimensions and relative position of the cavity and piston minimizes sound leakage and friction between them.

9:25

**4aEA5. Anechoic measurements of particle-velocity probes compared to pressure gradient and pressure microphones.** Wieslaw Woszczyk (CIRMMT, McGill Univ., 555 Sherbrooke St. West, Montreal, QC, Canada H3A 1E3, wieslaw@music.mcgill.ca), Masakazu Iwaki, Takehiro Sugimoto, Kazuho Ono (NHK Sci. & Tech. Res. Labs., Setagaya-ku, Tokyo 157-8510, Japan), and Hans-Elias de Bree (R&D Microflown Technologies)

Microflown probes are true figure-of-eight-pattern velocity microphones having extended response down to below the lowest audible frequencies, low noise, and high output. Unlike pressure-gradient microphones, velocity probes do not measure acoustic pressure at two points to derive a pressure gradient. When particle velocity is present, acoustical particle velocity sensors measure the temperature difference of the two closely spaced and heated platinum wire resistors, and quantify particle velocity from the temperature measurement. Microflown probes do not require a membrane and the associated mechanical vibration system. A number of anechoic measurements of velocity probes are compared to measurements of pressure-gradient and pressure microphones made under identical acoustical conditions at varying distances from a point source having a wide frequency response. Detailed measurements show specific response changes affected by the distance to the source, and focus on the importance of transducer calibration with respect to distance. Examples are given from field work using microflown probes to record acoustic response of rooms to test signals. The probe's cosine directional selectivity can be used to change the ratio between early reflections and the diffuse sound since only the  $\frac{1}{3}$  of the power in the diffuse sound field is measured with the particle velocity probe.

9:45

**4aEA6. Sensitivity change with practical use of electret condenser microphone.** Yoshinobu Yasuno (Panasonic Semiconductor Device Solutions Co., Ltd. 600, Saedo-cho, Tsuzuki-ku, Yokohama, 224-8539, Japan) and Kenzo Miura (Panasonic Mobile Commun. Eng. Co., Ltd., Yokohama, Japan)

Dr. Sessler and Dr. West invented the electret condenser microphone (ECM) in 1966. It has since been applied in various ways as a sound input device. The ECM has become an important component as a microphone for communications because of its stable sensitivity frequency characteristic. Materials and production methods have been improved continually up to the present. In particular, the ECM reliability is based on the electret's stability. For that reason, the electret surface charge decay is the main factor in ECM sensitivity degradation. This study analyzed the changes of an ECM preserved for 28 years in the laboratory and actually used for an outdoor interphone unit for 29 years. The change of diaphragm stiffness and electret surface voltage were compared with the evaluation result of a heat-acceleration test and verified. A degradation estimate of sensitivity change of ECM was performed. Regarding the life of the electret predicted in the report of former study [K. Miura and Y. Yasuno, J. Acoust. Soc. Jpn. (E) **18**(1), 29–35 (1997)], the validity was verified using actual data from this long-term observation.

10:20

**4aEA7. Development of a small size narrow directivity microphone.** Masakazu Iwaki, Kazuho Ono, Takehiro Sugimoto (NHK Sci. & Technol. Res. Labs., 1-10-11 Kinuta, Setagaya-ku, Tokyo, 157-8510, Japan), Takeshi Ishii, and Keishi Inamaga (Sanken Microphone Co., Ltd, 2-8-8 Ogikubo, Suginami-ku, Tokyo, 167-0051, Japan)

We developed a new microphone that has very sharp directivity even in the low-frequency band. In an ordinary environment of sound pick-up, the energy of background noise is distributed mainly in frequencies lower than 1000 Hz. In such frequencies, a typical cardioid microphone has the directivity pattern close to that of omni-cardioid. Consequently, it is difficult to pick up the objective sounds clearly from background noises. To suppress the noises with very small level, the directivity pattern should be also sharpened in the low-frequency band. In this report, we describe a new method to sharpen directivity for the low-frequency band. The method requires three microphone capsules. One capsule is the main microphone with a very short acoustic pipe. The others compose a second-order gradient microphone to cancel the signal that comes from behind the main microphone. A special feature of this microphone is to control a dip frequency of behind sensitivity without changing the frequency response of the front sensitivity.

10:40

**4aEA8. Two-wafer bulk-micromachined silicon microphones.** Jianmin Miao and Chee Wee Tan (Micromachines Ctr., School of Mech. and Aerosp. Eng., Nanyang Technol. Univ., 50 Nanyang Ave., Singapore 639798, mjmmiao@ntu.edu.sg)

A two-wafer concept is proposed for silicon microphone manufacturing by using bulk-micromachining and wafer bonding technologies. Acoustical holes of the backplate in one wafer are micromachined by deep reactive ion etching and the diaphragm on another wafer is created by wet-chemical etching. The two wafers are then bonded together to form silicon condenser microphones. In order to minimize the mechanical-thermal noise and increase the sensitivity within the required bandwidth, an analytical model based on Zuckerwar's equations has been developed to find the optimum location of the acoustical holes in the backplate of microphones. According to our study, this analytical modeling has shown excellent agreement between the simulated and measured results for the B&K MEMS microphone. Silicon condenser microphones have been further optimized in terms of the air gap, number and location of acoustical holes to achieve the best performance with a low polarization voltage, and easy fabrication for possible commercialization. Details of analytical modeling, fabrication, and measurement results will be presented.

11:00

**4aEA9. Infrasound calibration of measurement microphones.** Erling Frederiksen (Briel & Kjaer, Skodsborgvej 307, 2850 Naerum, Denmark, erlingfred@bksv.com)

Increasing interest for traceable infrasound measurements has caused the Consultative Committee for Acoustics, Ultrasound and Vibration (CCAUV) of BIPM to initiate a key comparison calibration project (CCAUV.A-K2) on pressure reciprocity calibration down to 2 Hz. Ten national metrology institutes, including the Danish Primary Laboratory of Acoustics (DPLA), take part in this project. In addition DPLA has started its own infrasound calibration project, which is described in this paper. The purposes of this project are verification of the CCAUV results and development of methods for calibration of general types of measurement microphone between 0.1 and 250 Hz. The project includes the design of an active comparison coupler, an experimental low-frequency reference microphone, and new methods for its frequency response calibration. One method applies an electrostatic actuator and requires a low-pressure measurement tank, while the other requires an additional microphone, whose design is closely related to that of the reference microphone that is to be calibrated. The overall calibration uncertainty ( $k=2$ ) for ordinary measurement microphones is estimated to less than 0.05 dB down to 1 Hz and less than 0.1 dB down to 0.1 Hz, if the reference is calibrated in the latter mentioned way, i.e., by the related microphones method.

### Contributed Papers

11:20

**4aEA10. Free-field calibration of 1/4 inch microphones for ultrasound by reciprocity technique.** Hironobu Takahashi, Takeshi Fujimori, Ryuzo Horiuchi, and Sojun Sato (Natl. Metrology Inst. of Japan, AIST, Tsukuba Central 3, 1-1-1 Umezono, Tsukuba, 305-8563 Japan)

Recently, equipment that radiates ultrasound radiation at frequencies far beyond the audible range is increasing in our environment. Such electronic equipment has switching regulators or inverter circuits, and many devices are unintended sources of ultrasound radiation. However, the effects of airborne ultrasound on human hearing and the human body have not been well investigated. To estimate the potential damage of airborne ultrasound radiation quantitatively, it is necessary to establish an acoustic standard for airborne ultrasound because the standard is a basis of acoustic measurement. With the intention of establishing a standard on airborne ultrasound, a free-field calibration system with an anechoic chamber was produced. The principle of free-field calibration techniques is introduced in this presentation. Type WS3 microphones (B&K 4939) were calibrated in the system to examine the calibration ability to be achieved. Results

showed that it can calibrate a microphone from 10 to 100 kHz with dispersion of less than 1 dB. In addition, the effects that were dependent on the uncertainty of the calibration are discussed based on those results.

11:35

**4aEA11. An environmentally robust silicon diaphragm microphone.** Norihiro Arimura, Juro Ohga (Shibaura Inst. of Technol., 3-7-5 Toyosu, Koto-ku, Tokyo, 135-8548, Japan), Norio Kimura, and Yoshinobu Yasuno (Panasonic Semiconductor Device Solutions Co., Ltd., Saedo-cho, Tsuzuki-ku, Yokohama, Japan)

Recently, many small microphones installed in cellular phones are the electret condenser microphones (ECMs) that contain an organic film diaphragm. Although FEP of fluorocarbon polymer is generally used as the electret material, silicon dioxide is also used. Recently ECMs have been made small and thin while maintaining the basic sound performance according to the market demand. In addition, environment tests and the reflow soldering mounting process have been adjusted to meet market requirements. On the other hand, the examination satisfied the demand as

the high temperature resistance was insufficient. This paper depicts an examination and a comparison of conventional ECM with the experimental model, a silicon diaphragm condenser microphone produced using the MEMS method. The silicon diaphragm satisfies high-temperature resis-

tance and stable temperature characteristics because of its very small coefficient of linear expansion and it is measured total harmonic distortion (THD) on high pressure sound. Finally, it will be able to be used in high temperature and high pressure sound conditions in the future.

FRIDAY MORNING, 1 DECEMBER 2006

IAO NEEDLE/AKAKA FALLS ROOM, 8:15 TO 11:15 A.M.

### Session 4aMU

## Musical Acoustics: Music Information and Communication

Bozena Kostek, Cochair

*Gdansk Univ. of Technology, Multimedia Systems Dept., Narutowicza 11- 12, 80-952 Gdansk, Poland*

Masuzo Yanagida, Cochair

*Doshisha Univ., Dept. of Information Science and Intelligent Systems, 1-3 Tatara-Miyakodani, Kyo-Tanabe, Kyoto 610-0321, Japan*

### Invited Papers

8:15

**4aMU1. Introduction of the Real World Computing music database.** Masataka Goto (Natl. Inst. of Adv. Industrial Sci. and Technol. (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan, m.goto@aist.go.jp)

This paper introduces the *RWC (Real World Computing) Music Database*, a copyright-cleared music database that is available to researchers as a common foundation for research. Shared databases are common in other research fields and have contributed importantly to progress in those fields. The field of music information processing, however, has lacked a common database of musical pieces and a large-scale database of musical instrument sounds. The RWC Music Database was therefore built in fiscal 2000 and 2001 as the world's first large-scale music database compiled specifically for research purposes. It contains six original collections: the *Popular Music Database* (100 pieces), *Royalty-Free Music Database* (15 pieces), *Classical Music Database* (50 pieces), *Jazz Music Database* (50 pieces), *Music Genre Database* (100 pieces), and *Musical Instrument Sound Database* (50 instruments). To address copyright issues, all 315 musical pieces were originally composed, arranged, or performed, and all instrumental sounds were originally recorded. The database has already been distributed to more than 200 research groups and is widely used. In addition, a continuous effort has been undertaken to manually annotate a set of music-scene descriptions for the musical pieces, called *AIST Annotation*, which consists of the beat structure, melody line, and chorus sections.

8:35

**4aMU2. Japanese traditional singing on the same lyrics.** Ichiro Nakayama (Osaka Univ. of Arts, 469, Higashiyama, Kanan-cho, Minami-Kawachi-gun, Osaka, 585-8555 Japan) and Masuzo Yanagida (Doshisha Univ., Kyo-Tanabe, 610-0321 Japan)

Described is a database of Japanese traditional singing together with supplementary recording of Bel Canto for comparative studies. Singing sounds and spoken speech by the same singers are recorded in pair to form the body of the database. This database covers most of genres of Japanese traditional singing, such as Shinto prayers, Buddhist prayers, Nor, Kyogen, Heikyoku, Sokyoku, Gidayu-bushi, Kabuki, Nagauta, Tokiwazu, Kiyomoto, Itchu-bushi, Shinnai, Kouta, Zokkyoku, Rokyoku, Shigin, Ryukyu-clasico, Goze-uta, etc. All the sounds were recorded in anechoic chambers belonging to local institutions, mainly in Osaka and Tokyo, asking 78 professional singers including 18 "Living National Treasures" to sing as informants. The most important point of this database is that an original lyric especially prepared for this recording is commonly used to make comparative studies easy. All the subjects are asked to sing the common lyrics in their own singing styles. Shown here are comparisons of formant shifts in vowels from ordinary speaking to singing for some singers, and comparison of temporal features of fundamental frequency between Japanese traditional singing and Western Bel Canto. [Work supported by the Academic Frontier Project, Doshisha University.]

8:55

**4aMU3. Computational intelligence approach to archival musical recordings.** Andrzej Czyzewski, Lukasz Litwic, and Przemyslaw Maziewski (Gdansk Univ. of Technol., Narutowicza 11/12, 80-952 Gdansk, Poland)

An algorithmic approach to wow defect estimation in archival musical recordings is presented. The wow estimation is based on the simultaneous analysis of many sinusoidal components, which are assumed to depict the defect. The rough determination of sinusoidal components in analyzed musical recording is performed by standard sinusoidal modeling procedures employing magnitude and phase spectra analysis. Since archival recordings tend to contain distorted tonal structure, the basic sinusoidal modeling approach is often found insufficient, resulting in audible distortions in the restored signal. It is found that the standard sinusoidal modeling approach is prone to errors, especially when strong frequency or amplitude variations of sinusoidal components occur. It may result in gaps or inappropriately matched components, leading to incorrect estimation of the wow distortion. Hence, some refinements to sinusoidal component analysis including interpolation and extrapolation of tonal components are proposed. As it was demonstrated in experi-

ments, due to the nonlinear nature of wow distortion, the enhancement of sinusoidal analysis can be performed by means of a neural network. The paper demonstrates implemented algorithms for parasite frequency modulation in archival recordings together with obtained results. [Work supported by the Commission of the European Communities, within the Integrated Project No. FP6-507336: PRESTOSPACE.]

9:15

**4aMU4. Music information retrieval seen from the communication technology perspective.** Bozena Kostek (Gdansk Univ. of Technol., Narutowicza 11/12, PL-80-952 Gdansk, Poland)

Music information retrieval (MIR) is a multidiscipline area. Within this domain one can see various approaches to musical instrument recognition, musical phrase classification, melody classification (e.g., query-by-humming systems), rhythm retrieval, high-level-based music retrieval such as looking for emotions in music or differences in expressiveness, and music search based on listeners' preferences. One may also find research that tries to correlate low-level descriptor analysis to high-level human perception. Researchers from musical acoustics, musicology, and music domains on one side, and communication technology on the other side, work together within this area. This may foster a framework for broader and deeper comprehension of contributions from all these disciplines and, in addition, translate the automated access to music information, gathered in various forms around the World Wide Web, as a fully understandable process to all participants regardless of their background. The semantic description is becoming a basis of the next Web generation. Several important concepts have been introduced recently by the researchers associated with the MIR community with regard to semantic data processing including techniques for computing with words. In this presentation some aspects related to MIR are briefly reviewed in the context of possible and actual applications of ontology-based approach to this domain.

9:35

**4aMU5. Accompaniment included song waveform retrieval based on framewise phoneme recognition.** Yuichi Yaguchi and Ryuichi Oka (Univ. of Aizu, Tsuruga, Ikkimachi, Aizuwakamatsu, Fukushima, 965-8580 Japan)

A novel approach is presented for a retrieval method that is useful for waveforms of songs with accompaniment. Audio signals of songs have some different acoustical characteristics from speech signals. Furthermore, the length per mora of signals is longer than that of speech. Therefore, the authors suggest a sound retrieval system for application to musical compositions, including songs, that extracts framewise acoustical characteristics and uses a retrieval method for absorbing phoneme length. First, the system prepares two sets of phoneme identification functions that have corresponding order, but for which phoneme sets belong to different environments of accompaniment-included or accompaniment-reduced. Next, musical compositions are put into database and the query song wave converts a waveform to a label sequence using framewise phoneme recognition derived by Bayesian estimation that applies each phoneme identification function according to whether it is accompaniment-included or not. Finally, the system extracts an interval area, such as query data, from a database using spotting recognition that is derived using continuous dynamic programming (CDP). Retrieval method results agree well with earlier results [Y. Yaguchi and R. Oka, AIRS2005, LNCS3689, 503–509 (2005)] that applied the same musical composition set without accompaniment.

9:55–10:10 Break

10:10

**4aMU6. Design of an impression-based music retrieval system.** Kimiko Ohta, Tadahiko Kumamoto, and Hitoshi Isahara (NICT, Keihanna Kyoto 619-0289, Japan, kimiko@nict.go.jp)

Impression-based music retrieval helps users to find musical pieces that suit their preferences, feelings, or mental states from among a huge volume of a music database. Users are asked to select one or more pairs of impression words from among multiple pairs that are presented by the system and to estimate each selected pair on a seven-step scale to input their impressions into the system. For instance, if they want to locate musical pieces that will create a happy impression, they should check the radio button "Happy" in the impression scale: Very happy–Happy–A little happy–Neutral–A little sad–Sad–Very sad. A pair of impression words with a seven-step scale is called an impression scale in this paper. The system calculates the distance between the impressions of each musical piece in a user-specified music database and the impressions that are input by the user. Subsequently, it selects candidate musical pieces to be presented as retrieval results. The impressions of musical pieces are expressed numerically by vectors that are generated from a musical piece's pitch, strength, and length of every tone using  $n$ -gram statistics.

### Contributed Papers

10:30

**4aMU7. Automatic discrimination between singing and speaking voices for a flexible music retrieval system.** Yasunori Ohishi, Masataka Goto, Katunobu Itou, and Kazuya Takeda (Grad. School of Information Sci., Nagoya Univ., Furo-cho 1, Chikusa-ku, Nagoya, Aichi, 464-8603, Japan, ohishi@sp.m.is.nagoya-u.ac.jp)

This paper describes a music retrieval system that enables a user to retrieve a song by two different methods: by singing its melody or by saying its title. To allow the user to use those methods seamlessly without changing a voice input mode, a method of automatically discriminating between singing and speaking voices is indispensable. We therefore first investigated measures that characterize differences between singing and

speaking voices. From subjective experiments, we found that human listeners discriminated between these two voices with 70% accuracy for 200-ms signals. These results showed that even short-term characteristics such as the spectral envelope represented as MFCC can be used as a discrimination cue, while the temporal structure is the most important cue when longer signals are given. According to these results, we then developed the automatic method of discriminating between singing and speaking voices by combining two measures: MFCC and an  $F0$  (voice pitch) contour. Experimental results with our method showed that 68.1% accuracy was obtained for 200-ms signals and 87.3% accuracy was obtained for 2-s signals. Based on this method, we finally built a music retrieval system that can accept both singing voices for the melody and speaking voices for the title.

**4aMU8. Various acoustical aspects of an Asian (South) Indian classical music concert.** M. G. Prasad (Dept. of Mech. Eng., Stevens Inst. of Technol., Hoboken, NJ 07030), V. K. Raman (Flautist, Germantown, MD 20874), and Rama Jagadishan (Edison, NJ 08820)

An Asian (South) Indian classical music concert is an integrated acoustical experience for both the audience and the player(s). A typical concert team, either vocal or instrumental, consists of a main vocalist (or an instrumentalist) accompanied by a violinist, up to three percussion instrument players, and a reference drone. The concert is comprised of many songs. Each song has two main parts, namely Alapana and Kriti. The Alapana is an elaboration of a raga (tune) and the Kriti refers to the lyrics of the song. The violinist actively follows and supports the main musician during the concert. The percussion player(s) are provided an opportunity to present a solo of their rhythmic skills. The players and the audience communicate emotionally and intellectually with each other. Elements such as aesthetics, rhythm, skill, and emotional aspects of the players are evaluated and appreciated by the audience. This talk will present various aspects of a concert that brings about an integrated and holistic experience for both the audience and the player(s). Some samples from live vocal and instrumental music concerts will be presented.

**4aMU9. Musical scales, signals, quantum mechanics.** Alpar Sevgen (Dept. of Phys., Bogazici Univ., Bebek 34342, Istanbul, Turkey)

Scales, being finite length signals, allow themselves to be treated algebraically: key signatures are related to the “ring” property of the scale labels; cyclically permuted scales and their mirror images have the same number of sharps and flats; and complementary scales (like major and pentatonic) have their sharp and flat numbers exchanged. A search for minimum principles to select among all possible scales those employed in music yields two possibilities: (a) minimize total number of accidentals and (b) minimize frequency fluctuations in a scale. Either of these minimum principles helps filter those scales employed in music from the universe of all scales, setting up very different criteria than the harmonic ratios used by musicians. The notes of the scales employed in music seem to prefer to stay as far apart from each other as possible. Operators that step through the multiplet members of scales with  $N$  semitones form a complete set of operators together with those that step through their eigenvectors. The mathematics reveals the discrete Fourier transformations (DFT) and is identical to finite state quantum mechanics of  $N$ -level Stern-Gerlach filters worked out by J. Schwinger.

FRIDAY MORNING, 1 DECEMBER 2006

MAUI ROOM, 7:30 A.M. TO 12:15 P.M.

### Session 4aNS

## Noise and Architectural Acoustics: Soundscapes and Cultural Perception I

Brigitte Schulte-Fortkamp, Cochair

*Technical Univ. Berlin, Inst. of Technical Acoustics, Secr TA 7, Einsteinufer 25, 10587 Berlin, Germany*

Bennett M. Brooks, Cochair

*Brooks Acoustics Corp., 27 Hartford Turnpike, Vernon, CT 06066*

### Invited Papers

7:30

**4aNS1. Soundscape in the old town of Naples: Signs of cultural identity.** Giovanni Brambilla (CNR Istituto di Acustica “O.M. Corbino” Via del Fosso del Cavaliere 100, 00133 Roma, Italy), Luigi Maffei, Leda De Gregorio, and Massimiliano Masullo (Second Univ. of Naples, 81031 Aversa (Ce), Italy)

Like all cities in Magna Grecia, the ancient Neapolis was built along three main parallel, tight, and straight streets called *decumani*. Since then and during the following centuries, commercial and handicraft activities, as well as social life, have been developed along these streets. The narrow ground rooms forced shopkeepers to occupy the main street to show their merchandise using vocal appeals to magnify their product, and handicrafts to work directly on the street (hammering, sawing, etc.). Music artists had their performance on the streets too. The soundscape in the area was a strong symbol of the Neapolitan cultural identity. Nowadays *decumani* have kept the main features of the past but some of these are overrun by road traffic. To investigate in which way the traffic noise has modified the soundscape perception and cultural identity, sound walks were registered during day and night time. A number of residents were interviewed and laboratory listening tests were carried out. Despite the congested urban environment and high sound levels, preliminary results have shown that part of the residential population is still able to identify the soundscape more related to Neapolitan historical identity.

7:50

**4aNS2. Soundscape design in public spaces: Concept, method, and practice.** Hisao Nakamura-Funaba and Shin-ichiro Iwamiya (Kyushu Univ., 4-9-1.Shiobaru, Minami-ku, Fukuoka 815-8540, Japan)

Soundscape design of public spaces necessitates consideration of whether a space has important meaning for the user. It is essential to imagine an ideal sound environment of that space. We designed an actual soundscape from the viewpoint of the environment, information, and decoration. In many cases, producing some silence in an environment becomes the first step of soundscape design. There is neither a special technology nor a technique when designing. It merely requires use of a general technology and techniques concerning sound according to location. A key point is knowledge of how to coordinate these technologies and techniques. For instance, silence was made first at the renewal project of Tokyo Tower observatory through cancellation of its commercial and call broadcasting functions and installation of sound-absorbing panels to the ceiling. Next, suitable and small sounds were added at various points. Guests can take time to enjoy viewing as a result.

## 8:10

**4aNS3. The daily rhythm of soundscape.** Brigitte Schulte-Fortkamp (TU-Berlin, Einsteinufer 25 TA 7, 10587 Germany, brigitte.schulte-fortkamp@tu-berlin.de) and Andr Fiebig (HEAD acoustics GmbH, 52134 Herzogenrath, Germany)

With respect to people's minds, soundscapes can be considered as dynamic systems characterized by the time-dependent occurrence of particular sound events embedded in specific environments. Therefore, an adequate evaluation of environmental noise will reflect the continually varying acoustical scenery and its specific perception. An acoustical diary shall provide information about the daily routine and subjectively perceived sound exposure of residents. It relies on cognitive and emotional aspects of perceiving and evaluating sounds. It gives insight into evaluating processes and their contextual parameters because of its spontaneous character. It includes and refers to long-term acoustic measurements. Simultaneous measurements for elaborate acoustical analyses will be taken outside the homes regarding the moments of essential sound events. The aim is to collect information about the daily rhythm regarding acoustic events, whereby the focus also should be placed on sound events that are well accepted to deepen the explication of data with respect to the analysis. Procedure and results will be discussed.

## 8:30

**4aNS4. Ecological explorations of soundscapes: From verbal analysis to experimental settings.** Daniele Dubois (CNRS, 11 rue de Lourmel, 75015 Paris, France) and Catherine Guastavino (McGill Univ., Montreal, QC H3A 1Y1, Canada)

Scientific studies rely on rigorous methods that must be adapted to the object of study. Besides integrating acoustic features, soundscapes as complex cognitive representations also have the properties of being global, meaningful, multimodal, and categorical. Investigating these specificities, new paradigms were developed involving linguistics and ecological psychology to complement the psychophysical approach: cognitive linguistic analyses of discourse to address semantic properties of soundscapes, and categorization tasks and distances from prototypes to investigate their cognitive organization. As a methodological consequence, experimental settings must be designed to ensure the ecological validity of the stimuli processing, (the "realism" evaluated from a psychological point of view, stimuli being processed as in a real-life situation). This point will be illustrated with perceptual evaluations of spatial auditory displays for soundscape reproduction. Data processing techniques should also take into consideration the intrinsic properties of the representations they account for. Examples of free-sorting tasks will be presented with measurements in terms of family resemblance of sets of properties defining categories rather than dimensional scales. New ways of coupling physical measurement and psychological evaluations will be presented in order to simulate or reproduce soundscapes in both a realistic and controlled manner for experimental purposes.

## 8:50

**4aNS5. Artificial neural network models of sound signals in urban open spaces.** Lei Yu and Jian Kang (School of Architecture, Sheffield Univ., Western Bank, Sheffield S10 2TN, UK)

Sound signals, known as foreground sounds, are important components of soundscape in urban open spaces. Previous studies in this area have shown that sound preferences are different according to social and demographic factors of individual users [W. Yang and J. Kang, *J. Urban Des.*, **10**, 69–88 (2005)]. This study develops artificial neural network (ANN) models of sound signals for architects at design stage, simulating subjective evaluation of sound signals. A database for ANN modeling has been established based on large-scale social surveys in European and Chinese cities. The ANN models have consequently been built, where individual's social and demographic factors, activities, and acoustic features of the space and sounds are used as input variables while the sound preference is defined as the output. Through the process of training and testing the ANN models, considerable convergences have been achieved, which means that the models can be applied as practical tools for architects to design sound signals in urban open spaces, taking the characteristics of potential users into account. Currently ANN models combining foreground and background sounds are being developed.

## 9:10

**4aNS6. Describing soundscape and its effects on people where soundscape is understood as an expansion of the concept of noise engineering.** Keiji Kawai (Grad. School of Sci. and Technol., Kumamoto Univ., 2-39-1 Kurokami, Kumamoto 860-8555, Japan, kawai@arch.kumamoto-u.ac.jp)

This study discusses how to describe sound environment and people in terms of "soundscape" as an expansion of the "noise engineering." In the framework of the conventional study field of noise evaluation, typically, sound environments are represented by loudness-based indices such as A-weighted sound pressure levels, and the impact of sound environments on people is represented by annoyance response or some physiological metrics. In the case of soundscape studies, however, the description should be expanded beyond what has been used in noise engineering. This matter has already been frequently discussed, but it doesn't seem that much consensus has been achieved concerning it yet. With respect to the effects of sound environment on people, since the concept of soundscape focuses on personal and social meanings of environmental sounds including the historical or aesthetic contexts, the effects are considered to be represented not by a singular concept such as comfortableness or quietness, but by multiple dimensions of emotional and aesthetic concepts. Also, descriptions of sound environment should include some qualitative aspects, such as what types of sounds can be heard at what extent. In this paper, the methodology to describe human-soundscape relationships is discussed through a review of related studies.

9:30

**4aNS7. The sound environmental education aided by automated bioacoustic identification in view of soundscape recognition.** Teruyo Oba (Natural History Museum & Inst., Chiba, 955-2 Aoba-cho, Chuo-ku, Chiba-shi, Chiba-ken 260-8682 Japan, oba@chiba-muse.or.jp)

From the 2003–2004 JST projects, where the automated bioacoustic identification device Kikimimi-Zukin was introduced to the nature observation and environmental studies, it was learned that the activities encouraged children to take notice of sounds, become aware of the sound environment, and gain an insight into the soundscape. Sounds are often riddles to us, and hearing is the process to find out the facts and causes. It is more important to let children obtain appropriate clues to differentiate sounds by hearing and thinking for themselves than give them an immediate answer. Program on the Strength of Hearing was formulated to challenge children and have them enjoy hearing to identify, sharing what they hear with others, and observing environment through sounds. Kikimimi-Zukin reinforced the program by a step-by-step guide through the hearing process of scanning, focusing, characterizing, associating with relevant factors, and judging the identity. The experience not only brought them confidence in hearing but incentive to study nature and environment. With Kikimimi-Zukin children collected recordings and relevant information. Using the sound database, the local singing map and three-dimensional sound map were prepared. They facilitated communication on the local sound environment among children and with adults, leading to realization of their inner soundscape.

9:50

**4aNS8. Acoustic environmental problems at temporary shelters for victims of the Mid-Niigata Earthquake.** Koji Nagahata, Norio Suzuki, Megumi Sakamoto, Fuminori Tanba (Fukushima Univ., Kanayagawa 1, Fukushima City, Fukushima, 960-1296, Japan, nagahata@sss.fukushima-u.ac.jp), Shin-ya Kaneko, and Tetsuhito Fukushima (Fukushima Medical Univ., Fukushima, 960-1295, Japan)

An earthquake on 23 October 2004 inflicted heavy damage on the Mid-Niigata district. The earthquake isolated Yamakoshi village; consequently, all the village residents were forced to evacuate to temporary shelters in neighboring Nagaoka city for 2 months. Two types of temporary shelters were used: gymnasiums, and buildings with large separated rooms similar to community centers. A questionnaire survey and interviews ( $N=95$ ) were conducted to elucidate problems of the living environment at the temporary shelters. This study analyzed acoustic environmental problems there. Noise-related problems were noted by 40 respondents (46.5%): they were the fifth most frequently cited environmental problems. Several serious complaints, e.g., general annoyance at the shelters and footsteps of refugees at night, were only indicated by respondents who had evacuated to the gymnasiums. However, some problems, e.g., the clamor of children, including crying babies and voices of other refugees, were indicated by respondents irrespective of the type of the shelters to which they had been evacuated. Therefore, buildings like community centers were more desirable for temporary shelters, at least from the perspective of noise problems.

10:10–10:30 Break

10:30

**4aNS9. The burden of cardiovascular diseases due to road traffic noise.** Wolfgang Babisch (Dept. of Environ. Hygiene, Federal Environ. Agency, Corrensplatz 1, 14195 Berlin, Germany, wolfgang.babisch@uba.de) and Rokho Kim (WHO/EURO Ctr. for Environment and Health, 53113 Bonn, Germany)

Epidemiological studies suggest a higher risk of cardiovascular diseases, including high blood pressure and myocardial infarction, in subjects chronically exposed to high levels of road or air traffic noise. A new meta-analysis was carried out to assess a dose-response curve, which can be used for a quantitative risk assessment and to estimate the burden of cardiovascular disease attributable to environmental noise in European regions. Noise exposure was grouped according to 5 dB(A)-categories for the daytime outdoor average A-weighted sound pressure level, ( $L_{day}$ , 16 h: 6–22 h), which was considered in most studies. Information on night-time exposure ( $L_{night}$ , 8 h: 22–6 h or 23–7 h) was seldom available. However, approximations can be made with respect to  $L_{den}$  according to the European directive on the assessment and management of environmental noise. The strongest evidence of an association between community noise and cardiovascular endpoints was found for ischaemic heart diseases, including myocardial infarction and road traffic noise. The disability-adjusted life years lost for ischemic heart disease attributable to transport noise were estimated conservatively, assuming the same exposure patterns across the countries with an impact fraction 3% in the western European countries.

10:50

**4aNS10. Soundscape, moderator effects, and economic implications.** Cay Hehner (Henry George School of Social Sci., 121 E. 30th St., New York, NY 10016, chehner.hengeoschool@att.net) and Brigitte Schulte-Fortkamp (TU-Berlin, Berlin, Germany)

Soundscape is considered with respect to moderator effects and the contribution of economics. It will be questioned whether soundscapes can work as a moderator concerning noise annoyance. As shown by the different investigations concerning soundscapes, a definition of the meaning of soundscapes is necessary. Evidently, the moderating effect of a given environment and its soundscape has to be discussed on three levels: (1) extension of factors that describe annoyance, (2) peculiar feature of burdensome noise contexts, and (3) discrepancies of the social and economic status of people living in areas where the rebuilding will change the quality of the area. It has to be determined and analyzed to what extent the Georgist method of resources taxation, as recently exemplified, e.g., in Alaska and in Wyoming, can be instrumental in funding soundscapes to moderate noise annoyance as it has been the case in funding free education and allowing the distribution of a citizen's dividend.

4a FRI. AM

**4aNS11. Socio-cultural soundscape concepts to support policies for managing the acoustic environment.** Michiko So Finegold, Lawrence S. Finegold (Finegold & So, Consultants, 1167 Bournemouth Court Centerville, OH 45459-2647, m-so@pb3.so-net.ne.jp), and Kozo Hiramatsu (Kyoto Univ., Sakyou-ku, Kyoto 606-8501 Japan)

In the past half-century, considerable effort has been invested in the academic, technological, and political arenas to achieve an adequate acoustic environment. Various national and international policy guidance documents have made reasonable progress in establishing a framework for a common approach to minimizing environmental noise, such as documents from various national Environmental Protection Agencies, the World Health Organization, and the European Union. Although these documents have provided useful information for global application, they only address minimizing the negative side of the acoustic environment (i.e., noise), they focus primarily on acoustics issues at the national or international level, and they still have not adequately considered implementation issues related to socio-cultural differences. To deal with the practical problems that exist in the acoustic environment in the context of different cultures, continuing research and new policy guidance are needed to address different local situations and in a variety of cultural contexts. The Soundscape approach has been developing tools for describing the acoustic environment at the local level to address both the positive and negative aspects of the acoustic environment. In this paper, the evolving interdisciplinary aspect of the Socio-Cultural Soundscape will be discussed and key topics for future work will be recommended.

### *Contributed Papers*

11:30

**4aNS12. Initial steps for the determination of environmental noise quality—The perception-related evaluation of traffic noise.** Klaus Genuit, Sandro Guidati, Sebastian Rossberg, and Andr Fiebig (HEAD Acoust. GmbH, Ebertstrasse 30a, 52134 Herzogenrath, Germany, klaus.genuit@head-acoustics.de)

Directives call for actions against noise pollution and noise annoyance. But, how do we eliminate harmful effects including annoyance due to the exposure of environmental noise without understanding the perception and evaluation of environmental noise? How do we preserve environmental noise quality where it is good (Directive 2002/49/EC) without identifying descriptors for noise quality? Various soundscape approaches based on different methodologies have been developed in the past. But, the measurement procedures must be realizable without much effort in order to achieve acceptance from legislation. Therefore, procedures have to be developed that capture the complexity of human hearing, on the one hand, and are feasible with respect to economy and time, on the other hand. The European project Quiet City (6FP PL516420) is dealing with, among other aspects, vehicle pass-by noise as a typical environmental noise source and its evaluation. Results of the analysis based on subjective assessments and psychoacoustic analyses carried out with respect to the development of an annoyance index will be presented and discussed. Such an index will

provide valuable information for effective improvement of noise quality. The final aim is to develop a descriptor valid for complete traffic noise scenarios predicting environmental noise quality adequately.

11:45

**4aNS13. When objective permissible noise limits of a municipal planning process and a subjective noise ordinance conflict.** Marlund Hale (Adv. Eng. Acoust., 663 Bristol Ave., Simi Valley, CA 93065, mehale@aol.com)

In most communities, proposed new building projects are required to conform with planning, community development, zoning, and/or building and safety specifications and standards. In cases of allowable noise exposure and noise limits, where certain of these requirements are quite specific while others are purposefully vague, conflicts between residential and commercial neighbors can lead to extreme disagreement and needless litigation. This paper describes a recent situation occurring in an upscale beach community, the resulting conflict over existing noise sources that comply with the limits of the city planning and permitting process, and the interesting findings of the court following the civil and criminal litigation that followed. Some suggestions are given to avoid these conflicting policy situations.

### *Contributed Poster Paper*

Poster paper 4aNS14 will be on display from 7:30 a.m. to 12:15 p.m. The author will be at the poster from 12:00 noon to 12:15 p.m.

**4aNS14. The complexity of environmental sound as a function of seasonal variation.** Hideo Shibayama (3-7-5 Koutou-ku Tokyo, 135-8548, Japan, sibayama@sic.shibaura-it.ac.jp)

Residential land is performed for a surrounding area of a suburban local city. As a result of urbanization, an area of rich natural environments became narrow. For the animals and plants for whom a river and a forest are necessary, it becomes difficult to live. Environmental sound produced by the tiny insects in this area is changing from year to year. Catching the

conditions for the developmental observations and environmental preservation in natural environments, we continue to measure the environmental sound as the time-series data. We estimate the complexity for these waveforms of the measured environmental sound in the season when insects chirp and do not chirp. For estimation of the complexity, we evaluate by the fractal dimension of the environmental sound. Environmental sound in early autumn is mainly generated by insects in the grass and on the trees. And, the fractal dimension for the sound waveforms of chirping of insects is up to 1.8.

## Session 4aPA

## Physical Acoustics and Biomedical Ultrasound/Bioresponse to Vibration: Sound Propagation in Inhomogeneous Media I

James G. Miller, Cochair

*Washington Univ., Dept. of Physics, 1 Brookings Dr., St. Louis, MO 63130*

Mami Matsukawa, Cochair

*Doshisha Univ., Lab. of Ultrasonic Electronics, Kyoto 610-0321, Japan*

Chair's Introduction—7:30

*Invited Papers*

7:35

**4aPA1. *In vivo* measurement of mass density and elasticity of cancellous bone using acoustic parameters for fast and slow waves.** Takahiko Otani (Faculty of Eng., Doshisha Univ., Kyotanabe 610-0321 Japan)

Cancellous bone (spongy bone) is comprised of a porous network of numerous trabecular elements with soft tissue in the pore space. The effect of decreasing bone density, namely a symptom of osteoporosis, is greater for cancellous bone than for dense cortical bone (compact bone). Two longitudinal waves, the fast and slow waves, are clearly observed in cancellous bone, which correspond to "waves of the first and second kinds" as predicted by Biot's theory. According to experimental and theoretical studies, the propagation speed of the fast wave increases with bone density and that of the slow wave is almost constant. Experimental results show that the fast wave amplitude increases proportionally and the slow wave amplitude decreases inversely with bone density. However, the attenuation constant of the fast wave is almost independent of bone density and the attenuation constant of the slow wave increases with bone density. The *in vivo* ultrasonic wave propagation path is composed of soft tissue, cortical bone, and cancellous bone and is modeled to specify the causality between ultrasonic wave parameters and bone mass density of cancellous bone. Then, mass density and elasticity are quantitatively formulated and estimated.

7:55

**4aPA2. Is ultrasound appropriate to measure bone quality factors?** Pascal Laugier (Univ. Pierre et Marie Curie., UMR CNRS 7623, 15 rue de l'Ecole de medecine, 7506 Paris, France, laugier@lip.bhdc.jussieu.fr)

Theoretical considerations support the concept that quantitative ultrasound variables measured in transmission are mainly determined by bone microstructure and material properties. All these properties are attributes of bone other than its bone mineral density (BMD) that may contribute to its quality and thus to strength or fragility. However, the limitations of this approach for a BMD-independent characterization of bone quality, long questioned, have become indisputable. Such considerations have prompted a research aiming at the development of new methods capable of measuring bone quality factors. New ultrasonic approaches are being investigated that use ultrasonic backscatter, guided waves, or nonlinear acoustics for studying bone microstructure or microdamage. These approaches, combined with sophisticated theoretical models or powerful computational tools, are advancing ideas regarding ultrasonic assessment of bone quality, which is not satisfactorily measured by x-ray techniques.

8:15

**4aPA3. Scanning acoustic microscopy studies of cortical and trabecular bone in the femur and mandible.** J. Lawrence Katz, Paulette Spence, Yong Wang (Univ. of Missouri-Kansas City, 650 E. 25th St., Kansas City, MO 64108, katzjl@umkc.edu), Anil Misra, Orestes Marangos (Univ. of Missouri-Kansas City, Kansas City, MO 64110), Dov Hazony (Case Western Reserve Univ., Cleveland, OH 44106), and Tsutomu Nomura (Niigata Univ. Grad. School of Medical and Dental Sci., Niigata, Japan)

Scanning acoustic microscopy (SAM) has been used to study the micromechanical properties of cortical and trabecular bone in both the human femur and mandible. SAM images vary in gray level reflecting the variations in reflectivity of the material under investigation. The reflection coefficient,  $r = (Z2 - Z1)/(Z2 + Z1)$ , where the acoustic impedance (AI),  $Z = \rho v$ ,  $\rho$  is the materials local density and  $v$  is the speed of sound at the focal point;  $Z2$  represents the AI of the material,  $Z1$  that of the fluid coupling the acoustic wave from the lens to the material. Femoral cortical bone consists of haversian systems (secondary osteons) and interstitial lamellae, both of which show systematic variations of high and low AI from lamella to lamella. The lamellar components defining the edges of trabecular cortical bone exhibit the same lamellar variations as seen in cortical bone. Mandibular bone, while oriented perpendicular to the direction of gravitational attraction, exhibits the same cortical and trabecular structural organizations as found in the femur. It also exhibits the same systematic alternations in lamellar AI as found in femoral bone. Both femoral and mandibular cortical bone have transverse isotropic symmetry. Thus, modeling elastic properties requires only five independent measurements.

**4aPA4. The interaction between ultrasound and human cancellous bone.** Keith Wear (US Food and Drug Administration, 12720 Twinbrook Pkwy., Rockville, MD 20852, keith.wear@fda.hhs.gov)

Attenuation is much greater in cancellous bone than in soft tissues, and varies approximately linearly with frequency between 300 kHz and 1.7 MHz. At diagnostic frequencies (300 to 700 kHz), sound speed is slightly faster in cancellous bone than in soft tissues. A linear-systems model can account for errors in through-transmission-based measurements of group velocity due to frequency-dependent attenuation and dispersion. The dependence of phase velocity on porosity may be predicted from theory of propagation in fluid-filled porous solids. The dependence of phase velocity on frequency (negative dispersion) can be explained using a stratified two-component model. At diagnostic frequencies, scattering varies as frequency to the  $n$ th power where  $3 > n > 3.5$ . This may be explained by a model that represents trabeculae as finite-length cylindrical scatterers.

**4aPA5. Dependence of phase velocity on porosity in cancellous bone: Application of the modified Biot-Attenborough model.** Suk Wang Yoon and Kang Il Lee (Dept. of Phys. and Inst. of Basic Sci., SungKyunKwan Univ., Suwon 440-746, Republic of Korea)

This study aims to apply the modified Biot-Attenborough (MBA) model to predict the dependence of phase velocity on porosity in cancellous bone. The MBA model predicted that the phase velocity decreases nonlinearly with porosity. The optimum values for input parameters of the MBA model, such as compressional speed  $c_m$  of solid bone and phase velocity parameter  $s_2$ , were determined by comparing the prediction with the previously published measurements in human calcaneus and bovine cancellous bones. The value of the phase velocity parameter  $s_2 = 1.23$  was obtained by curve fitting to the experimental data only for 53 human calcaneus samples with a compressional speed  $c_m = 2500$  m/s of solid bone. The root-mean-square error (rmse) of the curve fit was 15.3 m/s. The optimized value of  $s_2$  for all 75 cancellous bone samples (53 human and 22 bovine samples) was 1.42 with the rmse of 55 m/s. The latter fit was obtained by using  $c_m = 3200$  m/s. Although the MBA model relies on empirical parameters determined from the experimental data, it is expected that the model can be usefully employed as a practical tool in the field of clinical ultrasonic bone assessment.

**4aPA6. Simulation of fast and slow wave propagations through cancellous bone using three-dimensional elastic and Biot's trabecular models.** Atsushi Hosokawa (Dept. of Elect. & Comp. Eng., Akashi Natl. Coll. of Tech., 679-3 Nishioka, Uozumi, Akashi, 674-8501 Hyogo, Japan, hosokawa@akashi.ac.jp)

The propagation of ultrasonic pulse waves in cancellous (trabecular) bone was numerically simulated by using three-dimensional finite-difference time-domain (FDTD) methods. In previous research [A. Hosokawa, J. Acoust. Soc. Am. **118**, 1782–1789 (2005)], two two-dimensional FDTD models, the commonly used elastic FDTD model and an FDTD model based on Biot's theory for elastic wave propagation in an isotropic fluid-saturated porous medium, were used to simulate the fast and slow longitudinal waves propagating through cancellous bone in the direction parallel to the main trabecular orientation. In the present study, the extended three-dimensional viscoelastic and Biot's anisotropic models were developed to investigate the effect of trabecular structure on the fast and slow wave propagations. Using the viscoelastic model of the trabecular frame comprised of numerous pore spaces in the solid bone, the effect of the trabecular irregularity, that is the scattering effect, on both the fast and slow waves could be investigated. The effect of the anisotropic viscous resistance of the fluid in the trabecular pore spaces on the slow wave could be considered using Biot's anisotropic model.

### Contributed Papers

**4aPA7. Ultrasonic characteristics of *in vitro* human cancellous bone.** Isao Mano (OYO Electric Co., Ltd., Joyo 610-0101 Japan), Tadahito Yamamoto, Hiroshi Hagino, Ryota Teshima (Tottori Univ., Yonago 683-8503 Japan), Toshiyuki Tsujimoto (Horiba, Ltd., Kyoto 601-8510 Japan), and Takahiko Otani (Doshisha Univ., Kyotanabe 610-0321 Japan)

Cancellous bone is comprised of a connected network of trabeculae and is considered as an inhomogeneous and anisotropic acoustic medium. The fast and slow longitudinal waves are clearly observed when the ultrasonic wave propagates parallel to the direction of the trabeculae. The propagation speed of the fast wave increases with bone density and that of the slow wave is almost constant. The fast wave amplitude increases proportionally and the slow wave amplitude decreases inversely with bone density. Human *in vitro* femoral head was sectioned to 10-mm-thick slices perpendicularly to the femoral cervical axis. These cancellous bone samples were subjected to the ultrasonic measurement system LD-100 using a narrow focused beam. The propagation speed and the amplitude of the transmitted wave both for the fast and slow waves were measured at 1-mm intervals. The local bone density corresponding to the measured points was obtained using a microfocus x-ray CT system. Experimental results show that the propagation speeds and amplitudes for the fast and slow waves are characterized not only by the local bone density, but also by the local trabecular structure.

**4aPA8. The effect of hydroxyapatite crystallite orientation on ultrasonic velocity in bovine cortical bone.** Yu Yamato, Kaoru Yamazaki, Akira Nagano (Orthopaedic Surgery, Hamamatsu Univ. School of Medicine, 1-20-1 Hamamatsu Shizuoka 431-3192, Japan, yy14@hama-med.ac.jp), Hirofumi Mizukawa, Takahiko Yanagitani, and Mami Matsukawa (Doshisha Univ., Kyotanabe, Kyoto-fu 610-0321, Japan)

Cortical bone is recognized as a composite material of diverse elastic anisotropy, composed of hydroxyapatite (HAp) crystallite and type 1 collagen fiber. The aim of this study is to investigate the effects of HAp orientation on elastic anisotropy in bovine cortical bone. Eighty cubic samples were made from the cortical bone of two left bovine femurs. Longitudinal wave velocity in orthogonal three axes was measured using a conventional pulse echo system. For evaluating the orientation of HAp crystallite, x-ray diffraction profiles were obtained from three surfaces of each cubic sample. The preference of crystallites showed strong dependence on the direction of surface. The C-axis of crystallites was clearly preferred to the bone axial direction. The velocity in the axial direction was significantly correlated with the amounts of HAp crystallites aligning to the axial axis. The HAp orientation and velocity varied according to the microstructures of the samples. The samples with Haversian structure

showed larger preference of crystallites than plexiform samples in the axial direction. These results show clear effects of crystallites orientation on velocity.

10:05–10:20 Break

10:20

**4aPA9. Frequency variations of attenuation and velocity in cortical bone *in vitro*.** Magali Sasso, Guillaume Haiat (Universite Paris 12, Laboratoire de Mecanique Physique, UMR CNRS 7052 B2OA, 61, avenue du General de Gaulle, 94010 Creteil, France), Yu Yamato (Hamamatsu Univ. School of Medicine, Hamamatsu, Shizuoka, 431-3192, Japan), Salah Naili (Universite Paris 12, 94010 Creteil, France), and Mami Matsukawa (Doshisha Univ., Kyotanabe, Kyoto-fu 610-0321, Japan)

The development of ultrasonic characterization devices for cortical bone requires a better understanding of ultrasonic propagation in this heterogeneous medium. The aim of this work is to investigate the frequency dependence of the attenuation coefficient and of phase velocity and to relate them with bone microstructure and anatomical position. One hundred twenty parallelepipedic samples (4–11 mm side) have been cut from three bovine specimens and measured four times with repositioning in transmission with a pair of 8-MHz central frequency transducers. Phase velocity and BUA could be determined with acceptable precision: coefficients of variation of 0.8% and 13%, respectively. Velocity dispersion and BUA values are comprised between –13 and 40 m/s/MHz and 2 and 12 dB/MHz/cm, respectively. Negative dispersion values were measured (similarly to trabecular bone) for 2% of the measured samples. BUA values were found to be smaller in plexiform than in Haversian structure and higher for porotic structure. BUA values were found to be the greatest in the postero-lateral distal part and the smallest in the anterior-medial center part of the bone. The same tendency was found for velocity dispersion. Our results show the sensitivity of the frequency dependence of ultrasound to anatomical position and micro-architectural properties of bone.

10:35

**4aPA10. 3-D numerical simulations of wave propagation in trabecular bone predicts existence of the Biot fast compressional wave.** Guillaume Haiat (Universite Paris 12, Laboratoire de Mecanique Physique, UMR CNRS 7052 B2OA, 61, avenue du General de Gaulle, 94010 Creteil, France), Frederic Padilla, and Pascal Laugier (Universite Pierre et Marie Curie, 75006 Paris, France)

Trabecular bone is a poroelastic medium in which the propagation of two longitudinal waves (fast and slow) has been observed. The 3-D finite-difference time-domain simulations neglecting absorption coupled to 3-D microstructural models of 34 trabecular bone reconstructed from synchrotron radiation microtomography are shown to be suitable to predict both types of compressional wave in the three orthogonal directions. The influence of bone volume fraction (BV/TV) on the existence of the fast and slow waves was studied using a dedicated iterative image processing algorithm (dilation, erosion) in order to modify all 34 initial 3-D microstructures. An automatic criteria aiming at determining the existence of both wave modes was developed from the analysis of the transmitted signals in the frequency domain. For all samples, the fast wave disappears when bone volume fraction decreases. Both propagation modes were observed for BV/TV superior to a critical value for 2, 13, and 17 samples according to the direction of propagation. Above this critical value, the velocity of the fast (slow) wave increases (decreases) with BV/TV, consistent with Biot's theoretical predictions. This critical value decreases when the degree of anisotropy increases, showing coupling between structural anisotropy and the existence of the fast wave.

10:50

**4aPA11. 3-D numerical simulation of wave propagation in porous media: Influence of the microstructure and of the material properties of trabecular bone.** Guillaume Haiat (Universite Paris 12, Laboratoire de Mecanique Physique, UMR CNRS 7052 B2OA, 61, avenue du General de Gaulle, 94010 Creteil, France), Frederic Padilla, and Pascal Laugier (Universite Pierre et Marie Curie, 75006 Paris, France)

Finite-difference time domain simulations coupled to 3-D microstructural models of 30 trabecular bones reconstructed from synchrotron radiation microtomography were employed herein to compare and quantify the effects of bone volume fraction, microstructure, and material properties of trabecular bone on QUS parameters. Scenarios of trabecular thinning and thickening using an iterative dedicated algorithm allowed the estimation of the sensitivity of QUS parameters to bone volume fraction. The sensitivity to bone material properties was assessed by considering independent variations of density and stiffness. The effect of microstructure was qualitatively assessed by producing virtual bone specimens of identical bone volume fraction (13%). Both BUA and SOS show a strong correlation with BV/TV ( $r^2 = 0.94 p 10^{-4}$ ) and vary quasi-linearly with BV/TV at an approximate rate of 2 dB/cm/MHz and 11 m/s per % increase of BV/TV, respectively. Bone alterations caused by variation in BV/TV (BUA: 40 dB/cm.MHz; SOS: 200 m/s) is much more detrimental to QUS variables than that caused by alterations of material properties or diversity in microarchitecture (BUA: 7.8 dB/cm.MHz; SOS: 36 m/s). QUS variables are changed more dramatically by BV/TV than by changes in material properties or microstructural diversity. However, material properties and structure also appear to play a role.

11:05

**4aPA12. Singular value decomposition-based wave extraction algorithm for ultrasonic characterization of cortical bone *in vivo*.** Magali Sasso, Guillaume Haiat (Universite Paris 12, Laboratoire de Mecanique Physique, UMR CNRS 7052 B2OA, 61, avenue du General de Gaulle, 94010 Creteil, France), Maryline Talmant, Pascal Laugier (Universite Pierre et Marie Curie, 75006 Paris, France), and Salah Naili (Universite Paris 12, 94010 Creteil, France)

In the context of bone status assessment, the axial transmission technique allows ultrasonic evaluation of cortical bone using a multielement probe. The current processing uses the first arriving signal to evaluate the velocity while later contributions are potentially valuable and are not yet analyzed. However, all those contributions interfere, which disrupts the signal analysis. A novel ultrasonic wave extraction algorithm using a singular value decomposition method is proposed. This algorithm aims at characterizing a given energetic low-frequency (ELF) contribution observed *in vivo* at around 1 MHz. To evaluate the performances of the proposed algorithm, a simulated data set was constructed taking into account the influence of noise and of random interfering wavefront. The velocity of the ELF contribution is estimated on simulated datasets and compared to the input velocity. For a signal-to-noise ratio of 10 dB, the mean error associated with this method is 5.2%, to be compared with 34% with a classical signal analysis. The algorithm was also tested on real *in vivo* measurements. Results show the ability to accurately identify and possibly remove this wavefront contribution. Results are promising for multiple ultrasonic parameters evaluation from different wavefront contributions in our configuration.

11:20

**4aPA13. Direct evaluation of cancellous bone porosity using ultrasound.** Peiyong Liu, Matthew Lewis, and Peter Antich (Grad. Program of Biomed. Eng., Univ. of Texas Southwestern Medical Ctr. at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75390-9058)

Quantitative measurements of trabecular bone porosity would be a great advantage in diagnosis and prognosis of osteoporosis. This study focused on evaluation of the relationship between ultrasonic reflection and the density and porosity of cancellous bone. Theoretical simulation using MATLAB and Field II ultrasound simulator predicted a linear model of a material's porosity and parameters of the reflected ultrasonic signals. Ex-

perimentally, four plastic phantoms fabricated with different porosities were tested by a 5-MHz ultrasound transducer and the results agreed with simulations. Twelve specimens of bovine cancellous bone were measured. The porosities of these specimens were estimated by calculating the ratio of the mass in air to the wetted mass when they were immersed in water and all the air was removed from the pores. Among all the parameters, the peak value of the reflected ultrasound signal demonstrated highly significant linear correlations with porosity ( $R^2=0.911$ ) and density ( $R^2=0.866$ ). This encouraging result shows that this technique has the potential to be used to monitor porosity changes noninvasively for clinical purpose such as noninvasive assessment of osteoporotic fracture risk. [Work supported by Pak Foundation.]

11:35

**4aPA14. Transmission of ultrasound through bottlenose dolphin (*tursiops truncatus*) jaw and skull bone.** Michael D. Gray, James S. Martin, and Peter H. Rogers (Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405, michael.gray@gtri.gatech.edu)

Measurements of ultrasound transmission through jaw (pan bone) and skull (temporal fossa) samples from an Atlantic bottlenose dolphin were performed as part of an investigation of the feasibility of performing *in vivo* elastography on cetacean head tissues. The pan bone and temporal fossa are both relatively thin and smooth, and are therefore of interest for acoustic imaging of the underlying tissues of the ear and brain, respectively. Field scan data will be presented, showing the influence of the bone on the quality of the focus and overall pressure levels generated by a spherically focused single-element transducer. [Work supported by ONR.]

11:50

**4aPA15. Assessment of bone health by analyzing propagation parameters of various modes of acoustic waves.** Armen Sarvazyan, Vladimir Egorov, and Alexey Tatarinov (Artann Labs., Inc., 1457 Lower Ferry Rd., West Trenton, NJ 08618)

A method for assessment of bone based on comprehensive analysis of waveforms of ultrasound signals propagating in the bone is presented. A set of ultrasound propagation parameters, which are differentially sensitive to bone material properties, structure, and cortical thickness, are evaluated. The parameters include various features of different modes of acoustics

waves, such as bulk, surface, and guided ultrasonic waves in a wide range of carrier frequencies. Data processing algorithms were developed for obtaining axial profiles of waveform parameters. Such profiles are capable of revealing axial heterogeneity of long bones and spatially nonuniform pathological processes, such as osteoporosis. The examination procedure may include either one long bone in the skeleton, like the tibia, radius of the forearm, etc., or several bones in sequence to provide a more comprehensive assessment of the skeletal system. Specifically, for tibia assessment, a multi-parametric linear classifier based on a DEXA evaluation of skeleton conditions has been developed. Preliminary results of the pilot clinical studies involving 149 patients have demonstrated early stages of osteoporosis detection sensitivity of 80% and specificity of 67% based on DEXA data as the gold standard. [Work was supported by NIH and NASA grants.]

12:05

**4aPA16. Toward bone quality assessment: Interference of fast and slow wave modes with positive dispersion can account for the apparent negative dispersion.** James G. Miller, Karen Marutyan, and Mark R. Holland (Dept. of Phys., Washington Univ., St. Louis, MO 63130)

The goal of this work was to show that apparently negative dispersion in bone can arise from interference between fast wave and slow wave longitudinal modes, each of positive dispersion. Simulations were carried out using two approaches, one based on the Biot-Johnson model and the second independent of that model. The resulting propagating fast wave and slow wave modes accumulated phase and suffered loss with distance traveled. Results of both types of simulation served as the input into a phase and magnitude spectroscopy algorithm (previously validated with experimental data) in order to determine the phase velocity as a function of frequency. Results based on both methods of simulation were mutually consistent. Depending upon the relative magnitudes and time relationships between the fast and slow wave modes, the apparent phase velocities as functions of frequency demonstrated either negative or positive dispersions. These results appear to account for measurements from many laboratories that report that the phase velocity of ultrasonic waves propagating in cancellous bone decreases with increasing frequency (negative dispersion) in about 90% of specimens but increases with frequency in about 10%. [Work supported in part by Grant NIH R37HL40302.]

**Session 4aPP****Psychological and Physiological Acoustics and ASA Committee on Standards: New Insights on Loudness and Hearing Thresholds**

Rhona P. Hellman, Cochair

*Northeastern Univ., Dept. of Speech-Language Pathology and Audiology, Boston, MA 02115*

Yôiti Suzuki, Cochair

*Tohoku Univ., Research Inst. of Electrical Communication, Katarahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan***Chair's Introduction—8:30*****Invited Papers*****8:35****4aPP1. Threshold of hearing for pure tones between 16 and 30 kHz.** Kaoru Ashihara (AIST, Tsukuba Central, 6 Tsukuba, Ibaraki 305-8566 Japan, ashihara-k@aist.go.jp)

Hearing thresholds for pure tones were obtained at 2-kHz intervals between 16 and 30 kHz in an anechoic chamber. Measured 50 cm from the sound source, the maximum presentation sound pressure level ranged from 105 to 112 dB depending on the frequency. To prevent the listener from detecting the quantization noise or subharmonic distortions at low frequencies, a pink noise was added as a masker. Using a 3-down 1-up transformed up-down method, thresholds were obtained at 26 kHz for 10 out of 32 ears. Even at 28 kHz threshold values were obtained for 3 ears, but none were observed for a tone at 30 kHz. Above 24 kHz, the thresholds always exceeded 90 dB SPL. Between 16 and 20 kHz thresholds increased abruptly, whereas above 20 kHz the threshold increase was more gradual.

**8:55****4aPP2. Do we have better hearing sensitivity than people in the past?** Kenji Kurakata and Tazu Mizunami (Natl. Inst. of Adv. Industrial Sci. and Technol. (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki, 305-8566 Japan, kurakata-k@aist.go.jp)

Our hearing sensitivity to tones of various frequencies declines progressively as we become older. ISO 7029 describes a method for calculating expected hearing threshold values as a function of age. However, more than 30 years have passed since the ISO standard source data were published. An earlier paper of the present authors [K. Kurakata and T. Mizunami, *Acoust. Sci. Technol.* **26**, 381–383 (2005)] compared hearing threshold data of Japanese people in recent years to the ISO standard values. The results of that comparison showed that the age-related sensitivity decrease of Japanese people was smaller, on average, than those described in the standard. A large discrepancy was apparent at 4000 and 8000 Hz: more than 10 dB for older males. In response to that inference, the ISO/TC43/WG1 “threshold of hearing” initiated a project in 2005 to explore the possibility of updating ISO 7029. This paper presents a summary of those comparison results of audiometric data and the work in WG1 for revising the standard.

**9:15****4aPP3. Reliability and frequency specificity of auditory steady-state response.** Masaru Aoyagi, Tomoo Watanabe, Tsukasa Ito, and Yasuhiro Abe (Dept. of Otolaryngol., Yamagata Univ. School of Medicine, 2-2-2 Iida-Nishi, Yamagata, 990-9585, Japan)

The reliability and frequency specificity of 80-Hz auditory steady-state response (80-Hz ASSR) elicited by sinusoidally amplitude-modulated (SAM tones) tones and detected by phase coherence were evaluated as a measure of the hearing level in young children. The 80-Hz ASSR at a carrier frequency of 1000 Hz was monitored in 169 ears of 125 hearing-impaired children and auditory brainstem response (ABR) elicited by tone pips was examined in 93 ears. Both responses were examined during sleep, and the thresholds were compared with the behavioral hearing threshold, which was determined by standard pure-tone audiometry or play audiometry. In 24 ears with various patterns of audiogram, 80-Hz ASSRs were examined at different carrier frequencies, and the threshold patterns were compared with the audiograms to investigate the frequency specificity of ASSR. The correlation coefficient between the threshold of 80-Hz ASSR and pure-tone threshold ( $r=0.863$ ) was higher than that for ABR ( $r=0.828$ ). The threshold patterns of 80-Hz ASSR clearly followed the corresponding audiogram patterns in all types of hearing impairment. These findings suggest that 80-Hz ASSR elicited by SAM tones and detected by phase coherence is a useful audiometric device for the determination of hearing level in a frequency-specific manner in children.

**4aPP4. Use of perceptual weights to test a model of loudness summation.** Lori J. Leibold and Walt Jesteadt (Boys Town Natl. Res. Hospital, 555 N 30th St., Omaha, NE 68131)

We recently examined the contribution of individual components of a multitone complex to judgments of overall loudness by computing the perceptual weight listeners assign to each component in a loudness-matching task [Leibold *et al.*, *J. Acoust. Soc. Am.* **117**, 2597 (2005)]. Stimuli were five-tone complexes centered on 1000 Hz, with six different logarithmic frequency spacings, ranging from 1.012 to 1.586. When all components fell within the same critical band, weights varied little across components. In contrast, the range of weights increased with increasing frequency separation, with greater weight given to the lowest and highest frequency components. Perceptual weights were largely in agreement with the Moore *et al.* loudness model [J. Audio Eng. Soc. **45**, 224–237 (1997)], except at the widest bandwidth. In the current study we further examined predictions of the loudness model, focusing on the widest frequency-spacing condition. Masked thresholds and jnds for intensity discrimination were measured for each component and were compared to weights. The model predicts more interaction in the widely spaced conditions than simple critical band models, but underestimates the true interactions in conditions where components are widely spaced. Central factors appear to influence loudness, masked thresholds, and intensity discrimination in these conditions. [Work supported by NIH/NIDCD.]

**4aPP5. Increased loudness effect at the absolute threshold of hearing.** Junji Yoshida, Masao Kasuga (Grad. School of Eng., Utsunomiya Univ., 7-1-2 Yoto, Utsunomiya-shi, Tochigi-ken, 321-8585, Japan), and Hiroshi Hasegawa (Utsunomiya Univ., Tochigi-ken, 321-8585, Japan)

This study investigated the effects of a previous sound on loudness at the absolute threshold of hearing. Change of the absolute threshold of hearing was measured when a pure tone preceded the test tone in the measurement of the threshold. The previous sound at 60 dB SPL was presented first in one ear, followed by the presentation of the test sound in either the contralateral or ipsilateral ear at an interval of 0.5 s. Both the previous and test sounds had the same frequency of 500 Hz, and the same duration of 3 s. The change of the threshold was obtained from the difference between the thresholds with and without the previous sound. The threshold was found to be decreased significantly by approximately 2 dB when the previous sound was presented in the contralateral ear. On the other hand, the threshold was only slightly changed when the previous sound was presented in the ipsilateral ear. These experiment results suggest that loudness was increased by perceiving the previous sound in the contralateral ear.

#### 10:15–10:25 Break

**4aPP6. Induced loudness reduction.** Michael Epstein (Auditory Modeling and Processing Lab., Inst. for Hearing, Speech & Lang., Dept. of Speech-Lang. Path. and Aud., Northeastern Univ., Boston, MA 02115) and Mary Florentine (Northeastern Univ., Boston, MA)

Induced loudness reduction (ILR) is a phenomenon by which a preceding higher-level tone (an inducer tone) reduces the loudness of a lower-level tone (a test tone). Under certain conditions, ILR can result in loudness reductions of 10 to 15 phons for pure tones. The strength of this effect depends on a number of parameters including: (1) the levels of both the inducer and test tones; (2) the frequency separation between the inducer and test tones; (3) the durations of the inducer and test tones; (4) the time separation between the inducer and test tones; (5) individual differences; and, possibly (6) the number of exposures to inducers. Because of the sensitivity of ILR to procedural conditions, it is quite important to consider the potential effects of ILR when considering any experimental design in which level varies. The understanding of ILR has also given insight into a number of unexplained discrepancies between data sets that were collected using different procedures. In fact, some of the variability known to affect loudness judgments may be due to ILR. [Work supported by NIH/NIDCD Grant R01DC02241.]

**4aPP7. Loudness growth in individual listeners with hearing loss.** Jeremy Marozeau and Mary Florentine (Commun. Res. Lab., Inst. for Hearing, Speech & Lang., Northeastern Univ., Boston, MA 02115)

Recent research indicates that there are large individual differences in how loudness grows with level for listeners with sensorineural hearing losses of primarily cochlear origin. Studies of loudness discomfort level suggest that loudness for most of these impaired listeners approaches the loudness experienced by the normal listeners at high levels. Loudness growth at low levels is still controversial. Although it is now well established that loudness at threshold is greater than zero, its exact value is unknown. If this value is the same for normal and impaired listeners, then the loudness growth for the impaired listeners must be steeper in order to catch up to normal at high levels. This phenomenon is called recruitment. If the loudness at threshold for impaired listeners is higher than that for normal listeners, then the impaired listeners will no longer be able to perceive sounds as soft. This phenomenon is called softness imperception. Results from two experiments suggest that: (1) individual differences are more important for impaired listeners than for normal listeners; (2) some impaired listeners seem to show recruitment, others softness imperception; and (3) averaging the results across the impaired listeners will mask these differences. [Work supported by NIH/NIDCD grant R01DC02241.]

**4aPP8. Growth of loudness in cochlear implant listening.** Robert L. Smith and Nicole Sanpetrino (Inst. for Sensory Res. and Dept. of Biomed. and Chemical Eng., Syracuse Univ., 621 Skytop Rd., Syracuse NY, 13244)

Cochlear implants (CIs) roughly mimic the transformation from sound frequency to cochlear place that occurs in acoustic hearing. However, CIs are relatively less capable of creating the intensive transformations that normal peripheral auditory processing provides. This is partly because CIs have small operating ranges on the order of 10:1 in electric current, compared to the 1 000 000:1 operating range for sound-pressure level (SPL) in acoustic hearing. Furthermore, loudness in acoustic hearing grows as a compressive power function of SPL. In contrast, loudness reportedly grows as a more expansive function of current for CI users, i.e., a power law with a large exponent or an exponential function. Our results, obtained using the minimally biased method of magnitude balance without an arbitrary standard, reveal a previously unreported range of shapes of CI loudness functions, going from linear to power laws with exponents of 5 or more. The shapes seem related in part to the etiology of deafness preceding cochlear implantation, although the shapes can vary with stimulating conditions within a subject. Furthermore, differential sensitivity to changes in current appears to be related to the shape of the corresponding loudness function. Implications for sound processing in electric and acoustic hearing will be discussed.

FRIDAY MORNING, 1 DECEMBER 2006

MOLOKAI ROOM, 8:00 A.M. TO 12:00 NOON

### Session 4aSC

#### Speech Communication: Perception (Poster Session)

Katsura Aoyama, Cochair

*Texas Tech. Univ., Health Science Ctr., Dept. of Speech-Language and Hearing Science, Lubbock, TX 79430-6073*

Masato Akagi, Cochair

*JAIST, School of Information Science, 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan*

#### Contributed Papers

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

**4aSC1. Functional load of segments and features.** Mafuyu Kitahara (School of Law, Waseda Univ., 1-6-1 Nishiwaseda, Shinjuku-Ku, Tokyo, Japan)

The present paper proposes a new measure of functional load for segments and features. In a nut shell, it is based on word frequencies and the number of minimal pairs in which the relevant segment/feature is crucial in distinction. For example, minimal pairs distinguished only by /t/ are most frequent in English while those distinguished by /k/ are most frequent in Japanese. As for functional load of features, single-feature contrasts and multiple-feature contrasts are incorporated in the calculation. In Japanese, [high] alone distinguishes the largest number of minimal pairs while [voice] distinguishes words most frequently in cooperation with other features. Word frequency and familiarity database for English and Japanese are used to observe the commonalities and differences in both languages with respect to the proposed measure of functional load. This line of analysis suggests a better account for a certain phonological process being more typical in one language but not in the other. Functional load can be thought of as a part of the top-down information from the lexicon, which interacts with the bottom-up perceptual information in the process of spoken word recognition. Not only the ease of articulation and perceptual salience but also the functional load drives phonological processes.

**4aSC2. Cortical representation of processing Japanese phonemic and phonetic contrasts.** Seiya Funatsu (The Prefectural Univ. of Hiroshima, 1-1-71 Ujinahigashi, Minami-ku, Hiroshima, 734-8558 Japan), Satoshi Imaizumi (The Prefectural Univ. of Hiroshima, Gakuen-machi Mihara, 723-0053 Japan), Akira Hashizume, and Kaoru Kurisu (Hiroshima Univ., Minami-ku Hiroshima, 734-8551 Japan)

This study investigated how Japanese speakers process phonemic and phonetic contrasts using voiced and devoiced vowel /u/ and /÷u/. During six oddball experiments, brain responses were measured using magnetoencephalography. Under the phonemic condition, a frequent stimulus /ch÷ita/ was contrasted with a deviant /ts÷uta/, and a frequent /ts÷uta/ with a deviant /ch÷ita/. Under the phonetic condition, a frequent /ts÷uta/ was contrasted with a deviant /tsuta/, and a frequent /tsuta/ with a deviant /ts÷uta/. Under the segment condition, vowel segments, /÷u/ and /u/, extracted from spoken words, were contrasted. The subjects were 13 native Japanese speakers. The equivalent current dipole moment (ECDM) was estimated from the mismatch field. Under the phonetic condition, the ECDM elicited by the voiced deviant was significantly larger than that elicited by the devoiced deviant in both hemispheres ( $p < 0.01$ ), while there were no significant deviant-related differences in ECDM under the phonemic condition in both hemispheres. Under the segment condition, the ECDM elicited by the voiced deviant and devoiced deviant did not differ significantly in either hemispheres. These results suggested that the

ECDM asymmetries between the voiced and the devoiced deviant observed under the phonetic condition did not originate from the acoustical difference itself, but from the phonetic environment.

**4aSC3. Evaluating a model to estimate breathiness in vowels.** Rahul Shrivastav, Arturo Camacho, and Sona Patel (Dept. of Commun. Sci. and Disord., Univ. of Florida, Gainesville, FL 32611)

The perception of breathiness in vowels is cued by changes in aspiration noise (AH) and the open quotient (OQ) [Klatt and Klatt, *J. Acoust. Soc. Am.* **87**(2), 820–857 (1990)]. A loudness model can be used to determine the extent to which AH masks the harmonic components in voice. The resulting partial loudness (PL) and loudness of AH (noise loudness; NL) have been shown to be good predictors of perceived breathiness [Shrivastav and Sapienza, *J. Acoust. Soc. Am.* **114**(1), 2218–2224 (2005)]. The levels of AH and OQ were systematically manipulated for ten synthetic vowels. Perceptual judgments of breathiness were obtained and regression functions to predict breathiness from NL/PL were derived. Results show breathiness to be a power function of NL/PL when NL/PL is above a certain threshold. This threshold appears to be affected by the stimulus pitch. A second experiment was conducted to determine if the resulting power function could be used to estimate breathiness in natural voices. The breathiness of novel stimuli, both natural and synthetic, was determined in a listening test. For comparison, breathiness for the same stimuli was also estimated using the power function obtained previously. Results show the extent to which findings can be generalized. [Research supported by NIH/R21DC006690.]

**4aSC4. Predicting vowel inventories: The dispersion-focalization theory revisited.** Roy Becker (Dept. of Linguist., Univ. of California Los Angeles, 3125 Campbell Hall, Los Angeles, CA 90095-1543)

A revision of the dispersion-focalization theory (DFT) [Schwartz *et al.*, *J. Phonetics* **25**, 233–253 (1997)] is presented. Like DFT, the current computational model incorporates the center of gravity effect (COG) of 3.5-Bark spectral integration, but it deviates from DFT in that the COG contributes to the actual values and reliability weights of the perceived formants of vowels. The COG is reinterpreted as a domain of acceleration towards formant merger: the percepts of formants less than 3.5 Barks apart are perturbed towards one another in a nonlinear yet continuous fashion and their weights are increased, but perceptual merger and weight maximization occur only when the acoustic distance is about 2 Bark. Like other dispersion-based models, inventories are evaluated predominantly according to the least dispersed vowel pair, where dispersion is measured as the weighted Euclidean distance between the vowels coordinates (the first two perceived formants). Yet in the current model the weights are determined dynamically, in a well-principled manner. This model improves existing models in predicting certain universal traits, such as series of front rounded vowels in large vowel inventories, as emergent properties of certain local maxima of the inventory dispersion evaluation function, without sacrificing predictive adequacy for smaller inventories.

**4aSC5. Matching fundamental and formant frequencies in vowels.** Peter F. Assmann (School of Behavioral and Brain Sci., Univ. of Texas at Dallas, Box 830688, Richardson, TX 75083), Terrance M. Nearey (Univ. of AB, Edmonton, AB, Canada T6E 2G2), and Derrick Chen (Univ. of Texas at Dallas, Richardson, TX 75083)

In natural speech, there is a moderate correlation between fundamental frequency (F0) and formant frequencies (FF) associated with differences in larynx and vocal tract size across talkers. This study asks whether listeners prefer combinations of mean F0 and mean FF that mirror the covariation of these properties. The stimuli were vowel triplets (/i/-/a/-/u/) spoken by two men and two women and subsequently processed by Kawahara's STRAIGHT vocoder. Experiment 1 included two continua, each containing 25 vowel triplets: one with the spectrum envelope (FF) scale

factor fixed at 1.0 (i.e., unmodified) and F0 varied over  $\pm 2$  oct, the other with F0 scale factor fixed at 1.0 and FF scale factors between 0.63 and 1.58. Listeners used a method of adjustment procedure to find the "best voice" in each set. For each continuum, best matches followed a unimodal distribution centered on the mean F0 or mean FF (F1, F2, F3) observed in measurements of vowels spoken by adult males and females. Experiment 2 showed comparable results when male vowels were scaled to the female range and vice versa. Overall the results suggest that listeners have an implicit awareness of the natural covariation of F0 and FF in human voices.

**4aSC6. Acoustic cues for distinguishing consonant sequences in Russian.** Lisa Davidson and Kevin Roon (Linguist. Dept., New York Univ., 719 Broadway, 4th Fl, New York, NY 10003, lisa.davidson@nyu.edu)

In Russian, the same consonant sequences are permitted in various environments. Consequently, the presence of a word boundary or reduced vowel can be phonologically contrastive (e.g. [zɔdɐvatʲ] "to assign," [zɔdɐvaʲ] "to turn in"), and both learners and experienced listeners likely rely on fine acoustic cues to discriminate the phonotactic structures they hear. In this study, the acoustic characteristics of consonant sequences are examined to establish which cues may distinguish (a) word-initial clusters (C1C2); (b) consonant-schwa-consonant sequences (C1VC2); and (c) sequences divided by a word boundary (C1#C2). For all three sequence types, native Russian speakers produced phrases containing three categories of target sequences: stop+consonant, fricative+consonant, and nasal+consonant. Results show no significant differences in C1 burst duration for initial stops, though a longer interconsonantal duration is a reliable cue to schwa presence in C1VC2. C2 is significantly longer in C#C than in other sequences. For clusters, when C1 is a stop, there are no significant differences in duration with other sequence types, but other C1's are significantly shorter. This suggests that articulatory overlap, which may lead to C1 shortening for fricative or nasal-initial clusters, is reduced in stop-initial clusters to ensure that the stop is released and recoverable. [Research supported by NSF.]

**4aSC7. Lipread me now, hear me better later: Crossmodal transfer of talker familiarity effects.** Kaayumari Sanchez, Lawrence D. Rosenblum, and Rachel M. Miller (Dept. of Psych., Univ. of California, Riverside, Riverside, CA)

There is evidence that for both auditory and visual speech perception (lipreading) familiarity with a talker facilitates speech recognition [Nygaard *et al.*, *Psychol. Sci.* **5**, 42 (1994); Yakel *et al.*, *Percept. Psychophys.* **62**, 1405 (2000)]. Explanations of these effects have concentrated on the retention of talker information specific to each of these modalities. It could be, however, that some amodal, talker-specific articulatory style information is retained to facilitate speech perception in both modalities. If this is true, then experience with a talker in one modality should facilitate perception of speech from that talker in the other modality. To test this prediction, subjects were given one hour of experience lipreading from a talker and were then asked to recover speech-in-noise from either this same talker or from a different talker. Results revealed that subjects who lipread and heard speech from the same talker performed better on the speech-in-noise task than did subjects who lipread from one talker and then heard speech from a different talker.

**4aSC8. Acoustic patterns of Japanese voiced velar stops.** James Dembowski and Katsura Aoyama (Dept. Speech-Lang. & Hearing Sci., Texas Tech Univ. Health Sci. Ctr., 3601 4th St., Lubbock, TX 79430-6073, james.dembowski@ttuhsc.edu)

This presentation focuses on Japanese voiced velar /g/. The phoneme /g/ in VCV contexts is said to be characterized by a distinctive wedge-shaped formant pattern in which F2 and F3 converge toward one fre-

quency in the transition from vowel to stop closure, and then diverge as the vocal tract opens from the stop release to the following vowel. This pattern was examined using acoustic and kinematic data from an x-ray microbeam database of Japanese speakers, which is comparable to the English language x-ray microbeam speech production database [Hashi *et al.*, J. Acoust. Soc. Am. **104**, 2426–2437 (1998)]. Japanese speakers produced the expected wedge-shaped formant pattern in isolated VCV nonsense syllables, but rarely, if ever, produced this pattern in connected speech. In contrast, English speakers more frequently produced the expected formant pattern in connected speech, though the pattern was less reliably present than in isolated VCVs and varied considerably within and across speakers. These observations highlight substantial differences between controlled laboratory speech and meaningful connected speech, as well as differences in the ways that phonemes are manifested by different linguistic communities. These data also potentially illuminate the relationship among phonetic, acoustic, and kinematic levels of speech production.

**4aSC9. Finding perceptual categories in multidimensional acoustic spaces.** Eric Oglesbee and Kenneth de Jong (Dept. of Linguist., Indiana Univ., Bloomington, IN 47405, eoglesbe@indiana.edu)

Examining phonetic categorization in multidimensional acoustic spaces poses a number of practical problems. The traditional method of forced identification of an entire stimulus space becomes prohibitive when the number and size of acoustic dimensions becomes increasingly large. In response to this, Iverson and Evans [ICPhS (2003)] proposed an adaptive tracking algorithm for finding best exemplars of vowels in a multidimensional acoustic space. Their algorithm converged on best exemplars in a relatively small number of trials; however, the search method took advantage of special properties of the vowel space in order to achieve rapid convergence. In this paper, a more general multidimensional search algorithm is proposed and analyzed for inherent biases. Then, using the algorithm, the phonetic categorization of /p/ and /b/ in a five-dimensional acoustic space by native speakers of English is tested. Results showed that (a) there were no substantial long-term biases in the search method and (b) the algorithm appeared to identify important acoustic dimensions in the identification of /p/ and /b/ using relatively few trials. [Work supported by NSF BCS-04406540.]

**4aSC10. On the perception of epenthetic stops in American English.** Amalia Arvaniti, Ryan Shosted, and Cynthia Kilpatrick (Dept. of Linguist., UCSD, 9500 Gilman Dr., La Jolla, CA 92093, amalia@ling.ucsd.edu)

This study examines the perception of epenthetic stops in American English. Stimuli were created from *prince*, *prints*, *mince*, *mints*, *quince*, and *quints* by removing all traces of /t/ and splicing in 0–72 ms of silence, in 12-ms steps, with or without a following burst. Subjects saw the minimal pairs on screen and selected the word they heard. It was hypothesized that stimuli with bursts and longer closure would result in more *t* responses (*prince* identified as *prints*) and that frequent words (*prince/prints*) would be more difficult to distinguish than infrequent words (*quince/quints*), as our production results suggest that frequent pairs are more likely to be produced similarly. Results from 19 subjects show shorter response times with longer silence intervals, but no effect of burst or stimulus identity. Similarly, stimuli with bursts were not identified as *nts* words more frequently than those without. Generally, stimuli originating from *nts* words were more likely to be identified as such if they contained a burst, while the opposite was true for stimuli from *nce* words.

Finally, frequent words were less likely to be correctly identified than infrequent words, suggesting that /t/ epenthesis is not as widespread throughout the lexicon as generally believed.

**4aSC11. Phonetic alignment to visual speech.** Rachel M. Miller, Lawrence D. Rosenblum, and Kaoyumari Sanchez (Dept. of Psych., Univ. of California, Riverside, Riverside, CA 92521)

Talkers are known to produce allophonic variation based, in part, on the speech of the person with whom they are talking. This subtle imitation, or phonetic alignment, occurs during live conversation and when a talker is asked to shadow recorded words [e.g., Shockley, *et al.*, Percept. Psychophys. **66**, 422 (2004)]. What is yet to be determined is the nature of the information to which talkers align. To examine whether this information is restricted to the acoustic modality, experiments were conducted to test if talkers align to visual speech (lipread) information. Normal-hearing subjects were asked to watch an actor silently utter words, and to identify these words by saying them out loud as quickly as possible. These shadowed responses were audio recorded and naive raters compared these responses to the actors auditory words (which had been recorded along with the actors visual tokens). Raters judged the shadowed words as sounding more like the actors words than did baseline words, which had been spoken by subjects before the shadowing task. These results show that phonetic alignment can be based on visual speech, suggesting that its informational basis is not restricted to the acoustic signal.

**4aSC12. New anthropomorphic talking robot—investigation of the three-dimensional articulation mechanism and improvement of the pitch range.** Kotaro Fukui, Yuma Ishikawa, Takashi Sawa, Eiji Shintaku (Dept. of Mech. Eng., Waseda Univ., 3-4-1 Ookubo, Shinjuku-ku, Tokyo, Japan), Masaaki Honda (Waseda Univ., Saitama, Japan), and Atsuo Takaniishi (Waseda Univ., Shinjuku-ku, Tokyo, Japan)

We developed a new three-dimensional talking robot WT-6 (Waseda Talker-No. 6), which generates speech sounds by mechanically simulating articulatory motions as well as aeroacoustic phenomenon in the vocal tract. WT-6 consists of a 17-DOF mechanism (1-DOF lungs, 5-DOF vocal cords, 1-DOF jaws, 5-DOF tongue, and 4-DOF lips). It has 3-D lips, tongue, jaw, and velum, which form the 3-D vocal tract structure. It also has an independent jaw opening/closing mechanism, which controls the relative tongue position in the vocal tract as well as the oral opening. The previous robot, which had a 2-D tongue [J. Acoust. Soc. Am. **117**, 2541 (2005)], was not enough to realize precise closure to produce humanlike consonants such as /s/ or /t/. The new tongue, which could be controlled to form the 3-D shape, makes it possible to produce more realistic vocal tract shape. The vocal cord model was also improved by adding a new pitch control mechanism pushing from the side of the vocal cords. The pitch range becomes broader than the previous robot, which is enough to reproduce normal human speech. Preliminary experimental results showed that synthesized speech quality improves for vowels /a/, /u/ and /o/. Some experimental results and video demonstration of the talking robot will be presented.

**4aSC13. The role of familiarity, semantic context, and amplitude-modulation on sentence intelligibility.** Tom Carrell (Univ. of Nebraska—Lincoln, Lincoln, NE 68583, tcarrell@unl.edu) and Dawna Lewis (Boys Town Natl. Res. Hospital, Omaha, NE 68131)

Amplitude modulation has been demonstrated to greatly improve the intelligibility of time-varying sinusoidal (TVS) sentences [Carrell and Opie, Percept. Psychophys. **52** (1992); Barker and Cooke, Speech Commun. **27** (1999); Hu and Wang, Proceedings of ICASSP-02 (2002)]. It has been argued that the improvement is due to a bottom-up process that causes the acoustically independent components of the sentences to be perceptually grouped for further analysis by the auditory system. A recent study [Shapley and Carrell, J. Acoust. Soc. Am. **118** (2005)] indicated that

semantic information did not influence intelligibility levels of TVS or modulated TVS sentences. In virtually all other studies in which speech was distorted or degraded its intelligibility was improved by appropriate semantic context [Miller, *et al.*, JEP 41 (1951)]. It is possible that listeners' unfamiliarity with TVS speech might account for the difference. With one exception every study that has employed this type of stimulus provided the listener with very few practice sentences [Lewis, AAA (2005)]. The present experiment manipulated listeners' familiarity with TVS sentences to test this notion. Participants were presented with high- and low-predictability TVS and modulated-TV S sentences. Familiarity had a large effect on perception and intelligibility. Interpretation of previous findings is reconsidered in this light.

**4aSC14. On the relation of apparent naturalness to phonetic perceptual resolution of consonant manner.** Robert E. Remez, Claire A. Landau, Daria F. Ferro, Judith Meer, and Kathryn Dubowski (Dept. of Psych., Barnard College, 3009 Broadway, New York, NY 10027)

How does the qualitative experience of speech influence phonetic perception? Our perceptual studies of consonant place and voicing have revealed a dichotomous relation between phonetic sensitivity and naturalness. Auditory quality and phonetic sensitivity sometimes co-vary, while in other conditions phonetic sensitivity is indifferent to huge variation in naturalness. New tests are reported extending the research to the dimension of manner, a contrast correlated with qualitatively distinct acoustic constituents in normal production. Speech synthesis techniques were used to create naturalness variants through (1) variation in the excitation of a synthetic voicing source and (2) variation in the bandwidth of the formant centers. Listeners calibrated the relative naturalness of items drawn from the test series, and the acuity of perceivers to the contrast between fricative and stop manner was estimated with cumulative  $d'$  across the series in identification tests. Combined with our prior findings, these new tests show how intelligibility and naturalness can be either perceptually orthogonal or contingent aspects of consonant dimensions, offering a tool to understand normative functions in speech perception. [Research supported by NIH (DC00308).]

**4aSC15. Effects of signal levels on vowel formant discrimination for normal-hearing listeners.** Chang Liu (Dept. of Commun. Sci. and Disord., Wichita State Univ., Wichita, KS 67260)

The goal of this study was to examine effects of signal levels on vowel formant discrimination. Thresholds of formant discrimination were measured for F1 and F2 of four vowels in isolated vowels and sentences at three intensity levels: 70, 85, and 100 dB SPL for normal-hearing listeners using a three-interval, two-alternative forced-choice procedure with a two-down, one-up tracking algorithm. Results showed that formant thresholds were significantly affected by formant frequency, linguistic context, and signal levels. Thresholds of formant discrimination were increased as formant frequency increased and linguistic context changed from isolated vowels to sentences. The signal level indicated a rollover effect in which formant thresholds at 85 dB SPL are lower than at either 70 or 100 dB SPL in both isolated vowels and sentences. This rollover level effect could be due to reduced frequency selectivity and reduction in active cochlear non-linearity at high signal levels for normal-hearing listeners. Excitation and loudness models will be explored to account for the level effect on formant discrimination of isolated vowels.

**4aSC16. Study on nonaudible murmur speaker verification using multiple session data.** Mariko Kojima, Hiromichi Kawanami, Hiroshi Saruwatari, Kiyohiro Shikano (Nara Inst. of Sci. and Technol. 8916-5 Takayama-cho Ikoma-shi, Nara, 630-0192 Japan), and Tomoko Matsui (The Inst. of Statistical Mathematics, Minato-ku, Tokyo, 106-8569 Japan)

A study on speaker verification with nonaudible murmur (NAM) segments using multiple session data was conducted. NAM is different from normal speech and is difficult for other people to catch. Therefore, a text-dependent verification strategy can be used in which each user utters her/

his own keyword phrase so that not only speaker-specific but also keyword-specific acoustic information is utilized. A special device called a NAM microphone worn on the surface of the skin below the mastoid bone is used to catch NAM because it is too low to be recorded using ordinary microphones. However, it is tolerant to exterior noise. This strategy is expected to yield relatively high performance. NAM segments, which consist of multiple short-term feature vectors, are used as input vectors to capture keyword-specific acoustic information well. To handle segments with a large number of dimensions, a support vector machine (SVM) is used. In experiments using NAM data uttered by 19 male and 10 female speakers in several different sessions, robustness against session-to-session data variation is examined. The effect of segment length is also investigated. The proposed approach achieves equal error rates of 0.04% (male) and 1.1% (female) when using 145-ms-long NAM segments.

**4aSC17. Sequential contrast or compensation for coarticulation?** John Kingston, Daniel Mash, Della Chambless, and Shigeto Kawahara (Linguist. Dept., Univ. of Massachusetts, Amherst, MA 01003-9274)

English listeners identify a stop ambiguous between /t/ and /k/ more often as "k" after /s/ than /sh/ (Mann and Repp, 1981). Judgments shift similarly after a fricative ambiguous between /s/ and /sh/ when its identity is predictable from a transitional probability bias but perhaps not from a lexical bias (Pitt and McQueen, 1998; cf. Samuel and Pitt, 2003). In replicating these experiments, we add a discrimination task to distinguish between the predictions of competing explanations for these findings: listeners respond "k" more often after /s/ because they compensate for the fronting of the stop expected from coarticulation with /s/ or because a stop with an F3 onset frequency midway between /t/'s high value and /k/'s low value sounds lower after the /s/'s high-frequency energy concentration. The second explanation predicts listeners will discriminate /s-k/ and /sh-t/ sequences better than /s-t/ and /sh-k/ sequences because sequential contrast exaggerates the spectral differences between /s-k/'s high-low intervals and /sh-t/'s low-high intervals and distinguishes them more perceptually than /s-t/'s high-high intervals and /sh-k/'s low-low intervals. Compensation for coarticulation predicts no difference in discriminability between the two pairs because it does not exaggerate differences between the two intervals. [Work supported by NIH.]

**4aSC18. Acoustics and perception of coarticulation at a distance.** Karen Jesney, Kathryn Pruitt, and John Kingston (Ling. Dept., Univ. of Massachusetts, Amherst, MA 01003-9274)

CVC syllables were recorded from two speakers of American English, in which the initial and final stops ranged over /b,d,g/ and the vowel ranged over /i,e,E,ae,u,U,o,O,a/. F2 locus equations differed systematically as a function of the place of articulation of the other stop. These equation's slope and y intercepts were used to synthesize initial /g-b/ and /g-d/ continua in CVC syllables in which the final stop ranged over /b,d,g/ and the vowel over /e,o,a/, and the resulting stimuli were presented to listeners for identification. Listeners responded g significantly more often to both continua when the final stop was /d/ rather than /b/; the number of g responses fell between the /d/ and /b/ extremes for final /g/. This difference between final /d/ vs. /b/ is only observed when the intervening vowel is back /o,a/ and is actually reversed weakly when it is front /e/. Listeners also respond g significantly more often when the final stop is /g/ rather than /b/ and the vowel is /o,a/ but not [e]. Segments do coarticulate at a distance, listeners take this coarticulation into account, and perceptual adjustments depend on the segments through which the coarticulation is expressed. [Supported by NIH.]

**4aSC19. Diphones, lexical access, and the verbal transformation effect.** James A. Bashford, Jr., Richard M. Warren, and Peter W. Lenz (Dept. of Psych., Univ. of Wisconsin—Milwaukee, P.O. Box 413, Milwaukee, WI 53201-0413, bashford@uwm.edu)

When listeners are presented with a repeating verbal stimulus, adaptation occurs and perception of the stimulus is replaced by perception of a competitor. The present study examines the first of these verbal transformations reported by 180 listeners who were presented with lexical and nonlexical consonantvowel (CV) syllables that varied in frequency-weighted neighborhood density (FWND). These stimuli were produced by pairing the six English stop consonants with a set of three vowels. As expected, the majority of initial illusory forms (78%) were neighbors, differing from the stimulus by a single phoneme, and the proportion of lexical neighbors increased with stimulus FWND. Interestingly, FWND had opposite effects upon the lability of consonants and vowels: There was a strong positive correlation [ $r=0.79$ ,  $F(17)=26.2$ ,  $p<0.0001$ ] between FWND and the number of consonant transformations, and in contrast, there was a strong negative correlation [ $r=-0.78$ ,  $F(17)=24.9$ ,  $p<0.0001$ ] between FWND and the number of vowel transformations. The implications of these and other findings with these simple diphones will be discussed in relation to current activation-competition theories of spoken word recognition. [Work supported by NIH.]

**4aSC20. Acoustic analysis and perceptual evaluation of nasalized /g/ consonant in continuous Japanese.** Hisao Kuwabara (Teikyo Univ. of Sci. & Technol., Uenohara, Yamashi 409-0193, Japan)

It is well known that the /g/ consonant, a velar voiced plosive, in Japanese continuous speech is often nasalized unless it appears at the word-initial position. The nasalized /g/ consonant takes place in dialects mainly spoken in northern districts including the Tokyo area where the standard Japanese is spoken. However, the number of nasalized /g/ consonants is said to be decreasing year by year according to a survey. This paper deals with acoustic and perceptual analysis of this phenomenon. Test materials used in this experiment are read version of Japanese short sentences by NHK's (Japan Broadcasting Corporation) professional announcers. Each sentence includes at least one /g/ consonant that would likely be pronounced as nasalized. An evaluation test reveals that less than 60% of nasalization has been found to occur for /g/ consonants for which 100% nasalization had been observed decades ago. Acoustic analysis for nasalized and non-nasalized /g/ sounds has been performed mainly through waveform parameters. It has been found that the power ratio between consonant and vowel is the most effective parameter for distinguishing nasals from non-nasals, but it is highly speaker dependent.

**4aSC21. Production and perception of place of articulation errors.** Adrienne M. Stearns and Stefan A. Frisch (Univ. of South Florida, 4202 E. Fowler Ave., PCD1017, Tampa, FL 33620)

Using ultrasound to examine speech production is gaining popularity because of its portability and noninvasiveness. This study examines ultrasound recordings of speech errors. In experiment 1, ultrasound images of participants' tongues were recorded while they read tongue twisters. Onset stop closures were measured using the angle of the tongue blade and elevation of the tongue dorsum. Measurements of tongue twisters were compared to baseline production measures to examine the ways in which erroneous productions differ from normal productions. It was found that an error could create normal productions of the other category (categorical errors) or abnormal productions that fell outside the normal categories (gradient errors). Consonant productions extracted from experiment 1 were presented auditory-only to naive listeners in experiment 2 for identification of the onset consonant. Overwhelmingly, the participants heard normal productions and gradient error productions as the intended sound. Categorical error productions were judged to be different from the intended sound. The only effect of erroneous production on perception appears to be a slight increase in reaction time, which may suggest that error tokens are abnormal in some way not measured in this study.

**4aSC22. Role of linguistic experience on audio-visual perception of English fricatives in quiet and noise backgrounds.** Yue Wang, Haisheng Jiang, Chad Danyluck (Dept. of Linguist., Simon Fraser Univ., Burnaby, BC, V5A 1S6, Canada), and Dawn Behne (Norwegian Univ. of Sci. and Technol., Trondheim, Norway)

Previous research shows that for native perceivers, visual information enhances speech perception, especially when auditory distinctiveness decreases. This study examines how linguistic experience affects audio-visual (AV) perception of non-native (L2) speech. Native Canadian English perceivers and Mandarin perceivers with two levels of English exposure (early and late arrival in Canada) were presented with English fricative-initial syllables in a quiet and a caf-noise background in four ways: audio-only (A), visual-only (V), congruent AV, and incongruent AV. Identification results show that for all groups, performance was better in the congruent AV than A or V condition, and better in quiet than in caf-noise background. However, whereas Mandarin early arrivals approximate the native English patterns, the late arrivals showed poorer identification, more reliance on visual information, and greater audio-visual integration with the incongruent AV materials. These findings indicate that although non-natives were more attentive to visual information, they failed to use the linguistically significant L2 visual cues, suggesting language-specific AV processing. Nonetheless, these cues were adopted by the early arrivals who had more L2 exposure. Moreover, similarities across groups indicate possible perceptual universals involved. Together they point to an integrated network in speech processing across modalities and linguistic backgrounds. [Work supported by SSHRC.]

**4aSC23. Voicing of /h/ in the Texas Instrument MIT (TIMIT) database.** Laura Koenig (Haskins Labs, 300 George St., New Haven, CT 06511 and Long Island Univ.-Brooklyn)

Although English /h/ is traditionally described as voiceless, authors have long recognized that voiced allophones exist, especially in unstressed, intervocalic positions. In past work, we have suggested that fully voiced /h/ may be more common in men than women, but our subject population was limited in number and dialectal diversity. In this study, we use the TIMIT database to obtain measures of /h/ production in men and women speaking multiple dialects of American English. Our analysis focuses on the /h/ initiating the word "had" in a sentence produced by all speakers in the database: "She had your dark suit . . ." Each token of /h/ is coded for whether (a) the /h/ is deleted (i.e., auditorily imperceptible); and, if /h/ is present, whether (b) voicing continues unbroken and (c) there is visible aspiration noise in the speech signal. This analysis will provide baseline data on /h/ realization in a common sentence context. We will also carry out follow-up analyses on selected utterances to gain more insight into the effects of phonetic context, stress, and lexical type (e.g., content versus function word) on the characteristics of English /h/. [Work supported by NIH.]

**4aSC24. On distinctions between stops and similar-place weak fricatives.** James M. Hillenbrand (Speech Pathol. and Audiol., MS5355, Western Michigan Univ., Kalamazoo, MI 49008) and Robert A. Houde (Ctr. for Commun. Res., Rochester, NY 14623)

There is an extensive body of literature on the acoustic properties of both stops and fricatives. However, little attention has been paid to the acoustic features that might distinguish these two sound categories. This is unsurprising in the case of stops versus strong fricatives since these sounds are rarely confused. Stops and weak fricatives, on the other hand, are frequently confused [G.A. Miller and P.E. Nicely, *J. Acoust. Soc. Am.* 27, 338–352 (1955)]. The present study was undertaken in a search for acoustic features that might distinguish the stop/weak fricative pairs /b/-/v/ and /d/-/dh/ (i.e., /d/ vs voiced th). Speech material consisted of CV and VCV syllables spoken by five men and five women, using the vowels /a/, /i/, and /u/. A combination of two features reliably separated

the stops from weak fricatives: (1) intensity during the consonant occlusion interval (typically greater for the fricatives), and (2) the rate of increase in mid- and high-frequency energy (above 1 kHz) associated with consonant release (typically greater for the stops).

**4aSC25. Salience of virtual formants as a function of the frequency separation between spectral components.** Robert A. Fox, Ewa Jacewicz, Chung-Yun Chang, and Jason D. Fox (Speech Acoust. and Percpt. Labs, Dept. of Speech and Hearing Sci., Ohio State, 110 Pressey Hall, 1070 Carmack Rd., Columbus, OH 43210-1002)

The center-of-gravity (COG) hypothesis proposed by Chistovich and others for the perception of static vowels suggests that auditory spectral integration may occur when two or more formants fall within a 3.5 bark bandwidth. While several studies have examined the bandwidth limits of such integration, this study examines the extent to which spectral integration is uniform within this putative 3.5-bark range. We examine the perceptual salience of virtual formants produced by modifying the spectral COG of two closely spaced narrow-bandwidth resonances. Three different vowel series were created: [i-i], [ $\Lambda$ - $\alpha$ ] and [ə-ə]. A second set of vowels was then created in which one of the formants (F1 in [i-i], F2 in [ $\Lambda$ - $\alpha$ ] and F3 in [ə-ə]) was replaced by a virtual formant whose COG matched that of the formant that had been removed. The frequency separation between the two component resonances was then systematically varied between 1.5 and 3.5 barks and a single-interval 2AFC vowel identification task was used to obtain estimates of vowel quality for each series step. Results will be discussed in terms of whether the spectral integration effects within the 3.5 bark decline as the frequency separation between the resonance components increases. [Work supported by NIDCD R01DC00679-01A1.]

**4aSC26. Frequency effects in phoneme processing.** Danny R. Moates, Noah E. Watkins, Zinny S. Bond, and Verna Stockmal (Inst. for the Empirical Study of Lang., Ohio Univ., Athens, OH 45701, moates@ohio.edu)

Are phonological segments activated during word recognition in proportion to their frequency of use? Previous evidence for this hypothesis [Moates *et al.*, *Laboratory Phonology 7*, edited by Gussenhoven and Warner (Mouton de Gruyter, 2002)] used a word reconstruction task. The present study used an online task, the gating task, in which progressively longer fragments of a word are presented to listeners who must identify the word in as few gates as possible. High- and low-frequency segments were contrasted by presenting them in word pairs that differed only in those two segments, e.g., label versus cable, where /l/ is used more frequently than /k/ [M. K. Vitevich and P. A. Luce, *Behav. Res. Methods, Instrum. Comput.* **36**, 481–487 (2004)]. We constructed 30 pairs of two-syllable words for which the contrasting segments were at the first syllable onset, 30 more for the second syllable onset, and 30 more for the coda of a syllable. Identification judgments were gathered from 120 participants. *t* tests showed high-frequency segments to be identified at significantly earlier gates than their matched low-frequency segments for first onset and coda, but not second onset. These results offer further evidence for sub-lexical processes in spoken word recognition.

**4aSC27. The clustering coefficient of phonological neighborhoods influences spoken word recognition.** Michael Vitevitch (Dept. of Psych., Univ. of Kansas, 1415 Jayhawk Blvd., Lawrence, KS 66045-7556)

Neighborhood density refers to the number of words, or neighbors, that are phonologically related to a given word. For example, the words BAT, MAT, CUT, and CAN (among others) are considered phonological neighbors of the word CAT. In contrast, the clustering coefficient of the neighborhood refers to the proportion of phonological neighbors that are also neighbors of each other. Among the neighbors of CAT, the words BAT and MAT are neighbors of each other, but the words BAT and CAN are not neighbors of each other. Despite the stimulus words having the

same number of neighbors overall, the results of an auditory lexical decision task showed that words with a high clustering coefficient (i.e., most neighbors were also neighbors of each other) were responded to more quickly than words with a low clustering coefficient (i.e., few neighbors were also neighbors of each other). These results suggest that some aspects of phonological similarity (i.e., clustering coefficient) might facilitate lexical activation, whereas other aspects of phonological similarity (i.e., neighborhood density) influence a later, decision stage of processing characterized by competition among activated word-forms. [Work supported by NIH.]

**4aSC28. The influence of noise and reverberation on vowel recognition: Response time.** Magdalena Blaszkak and Leon Rutkowski (Div. of Rm. Acoust. and Psychoacoustics, Adam Mickiewicz Univ., Umultowska 85, 61-614 Poznan, Poland)

This study examines the perceptual effect of two types of noise and reverberation on vowel recognition. Multitalker babble and traffic noise (European Standard EN 1793–3) were generated simultaneously with Polish vowels /a, e, i, o, u, y/ in two different sound fields and an anechoic chamber. The experiment was performed under various conditions of signal-to-noise ratio (–9, –6, –3, 0, +3, no noise). A new procedure for listeners' selection based on the Bourdon's psychometrics test was proposed. The effects of noise and reverberation were quantified in terms of (a) vowel recognition scores for young normal-hearing listeners (YNH) and (b) *ease of listening* based on the time of response and subjective estimation of difficulty. Results of the experiment have shown that (a) the response time can be a good measure of the effect of noise and reverberation on the speech intelligibility in rooms, and (b) in this type of experiment, of great significance is the choice of the subjects based on the psychometric tests.

**4aSC29. Quantifying the benefits of sentence repetition on the intelligibility of speech in continuous and fluctuating noises.** Isabelle Mercille, Roxanne Larose, Christian Giguère, and Chantal Laroche (Univ. of Ottawa, 451 Smyth Rd., Ottawa, ON, Canada K1H 8M5)

Good verbal communication is essential to ensure safety in the workplace and social participation during daily activities. In many situations, speech comprehension is difficult due to hearing problems, the presence of noise, or other factors. As a result, listeners must often ask the speaker to repeat what was said in order to understand the complete message. However, there has been little research describing the exact benefits of this commonly used strategy. This study reports original data quantifying the effect of sentence repetition on speech intelligibility as a function of signal-to-noise ratio and noise type. Speech intelligibility data were collected using 18 normal-hearing individuals. The speech material consisted of the sentences from the Hearing In Noise Test (HINT) presented in modulated and unmodulated noises. Results show that repeating a sentence decreases the speech reception threshold (SRT), as expected, but also increases the slope of the intelligibility function. Repetition was also found to be more beneficial in modulated noises (decrease in SRT by 3.2 to 5.4 dB) than in the unmodulated noise (decrease in SRT by 2.0 dB). The findings of this study could be useful in a wider context to develop predictive tools to assess speech comprehension under various conditions.

**4aSC30. The effect of the spectral shape changes on voice perception.** Mika Ito, Bruce R. Gerratt, Norma Antonanzas-Barroso, and Jody Kreiman (Div. of Head/Neck Surgery, UCLA School of Medicine, 31-24 Rehab Ctr., Los Angeles, CA 90095-1794, jkreiman@ucla.edu)

Researchers have long known that the shape of the vocal source spectrum is an important determinant of vocal quality, but the details regarding the importance of individual spectral features remains unclear. Previous research indicates four spectral features, H1-H2, the spectral slope above 4 kHz, the slope from 1.5–2 kHz, and the slope from 2–4 kHz, account for

virtually all the variability in spectral shapes. The present study provides preliminary evidence about the perceptual importance of these four features. Four series of stimuli were synthesized for each spectral parameter, in which that parameter varied in small steps. Because the perceptual salience of source parameters depends on F0 and on the spectrum of the inharmonic part of the source, series differed in the sex of the speaker (male/female) and in the NSR (noise-free/very noisy). Listeners heard all possible pairs of voices within each series and were asked to determine whether stimuli were the same or different. We hypothesize that listeners sensitivity to H1-H2 and the slope of the spectrum from 1.5–2 kHz will be independent of noise, but that sensitivity to changes in the spectral shape above 2 kHz will depend on the amount of noise excitation present in the voice.

**4aSC31. The use of auditory and visual information in the perception of stress in speech.** James Harnsberger, Daniel Kahan, and Harry Hollien (Inst. for Adv. Study of the Commun. Proc., Univ. of Florida, Gainesville, FL 32611)

Prior work on the acoustic correlates of the perception of psychological stress in speech has suffered from the problem of quantifying and verifying the extent to which a speaker was under stress during articulation. Two experiments were conducted to address this issue. First, stressed and unstressed speech samples were elicited from 78 speakers of American English. Stressed samples were recorded by having subjects read a standard passage while under the threat of the administration of mild electric shock. Both visual and audio recordings were collected. Stress was quantified in terms of four measures: two physiological (pulse rate and galvanic skin response) and two self-report scales. Sentences from the 16 speakers showing the largest differences between the stressed and unstressed conditions were then presented in a paired comparison task to 90 naive listeners, 30 each in three conditions: (1) audio-only presentation of the stimuli, (2) visual-only presentation of the stimuli, and (3) audiovisual presentation of the stimuli. The results indicate that individual listeners are sensitive to stress cues in speech in all three conditions.

**4aSC32. Selective attention and perceptual learning of speech.** Alexander L. Francis (Dept. of Speech, Lang. & Hearing Sci., Purdue Univ., West Lafayette, IN 47907), Natalya Kaganovich, and Courtney Driscoll (Purdue Univ., West Lafayette, IN 47907)

Phonetic experience can change the perceptual distance between speech sounds, increasing both within-category similarity and between-category distinctiveness. Such warping of perceptual space is frequently characterized in terms of changes in selective attention: Listeners are assumed to attend more strongly to category-differentiating features while ignoring less relevant ones. However, the link between the distribution of selective attention and categorization-related differences in perceptual distance has not been empirically demonstrated. To explore this relationship, listeners were given 6 h of training to categorize sounds according to one of two acoustic features while ignoring the other. The features were voice onset time and onset  $f_0$ , which are normally correlated and can both serve as a cue to consonant voicing. Before and after training, listener's performance on a Garner selective attention task was compared with assessment of the perceptual distance between tokens. Results suggest that training can induce both warping of perceptual space and changes in the distribution of selective attention, but the two phenomena are not necessarily related. Results are consistent with a two-stage model of perceptual learning, involving both preattentive adjustment of acoustic cue weighting and higher-level changes in the distribution of selective attention between acoustic cues. [Work supported by NIH-NIDCD 1R03DC006811.]

**4aSC33. Investigation of interaction between speech perception and production using auditory feedback.** Masato Akagi, Jianwu Dang, Xugang Lu, and Taichi Uchiyamada (School of Information Sci., JAIST, 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan)

This study employed an auditory feedback paradigm with perturbed fed-back speech to investigate interaction between speech perception and production by measuring simultaneous fluctuations of speech production organs using the electromyographic (EMG) signals, articulatory movements, as well as spectral analyses, where the articulatory data were obtained by the electromagnetic articulographic (EMA) system. Chinese vowels pair [i]-[y] and Japanese vowels pairs [e]-[a], [e]-[i] and [e]-[u] were chosen as the experimental objects. When the speaker is maintaining the first vowel, the feedback sound is randomly changed from the first vowel to the second one in each pair by manipulating the first three formants. Spectral analysis showed that a clear compensation was seen in the first and second formants of the vowels. Analyses of EMG and EMA signals also showed muscle reactivation and tongue movements to compensate for the perturbations. Latency of the compensating response is about 150 ms to start and about 290 ms for maximum compensation from the onset of the perturbation. According to the measurements, it seems that in most cases the speaker attempts to compensate for the "error" caused by the auditory perturbation by a real-time monitoring, and the auditory feedback takes place simultaneously often during speech production.

**4aSC34. Cross-ear suppression of the verbal transformation effect: Tweaking an acoustic-phonetic level.** Peter W. Lenz, James A. Bashford, Jr., and Richard M. Warren (Dept. of Psych., Univ. of Wisconsin—Milwaukee, P.O. Box 413, Milwaukee, WI 53201-0413, plenz@uwm.edu)

A recorded word repeating over and over undergoes a succession of illusory changes. When two images of the same repeating word are presented dichotically, with a half-cycle delay preventing fusion, the two images of the word each undergo independent illusory transformations at a rate equivalent to that of a single image [Lenz *et al.*, *J. Acoust. Soc. Am.* **107**, 2857 (2000)]. However, with one phoneme difference (e.g., "dark" versus "dart"), transition rate is dramatically suppressed [Bashford *et al.*, *J. Acoust. Soc. Am.* **110**, 2658 (2001)]. Rates decrease with extent of feature mismatch at a single phoneme position (roughly 30% reduction with one feature mismatch and 45% with three). Rates also decrease with the number of mismatched phonemes (about 80% rate reduction with three out of four), suggesting a strong acoustic-phonetic basis for verbal transformation suppression. In contrast, semantic relations had no effect (e.g., transformations for "light" were suppressed equally by contralateral night and "might"). Dichotic competition appears to allow us to access and selectively influence a prelexical stage of linguistic analysis. [Work supported by NIH.]

**4aSC35. Perceptually balanced filter response for binaural dichotic presentation to reduce the effect of spectral masking.** Pandurangarao N. Kulkarni, Prem C. Pandey (Elec. Eng. Dept, Indian Inst. of Technol. Bombay, Powai Mumbai 400076, India, pcPandey@ee.iitb.ac.in), and Dakshayani S. Jangamashetti (Basaveshwar Eng. College Bagalkot, Bagalkot Karnataka 587102, India)

Earlier investigations show that the scheme of binaural dichotic presentation with spectral splitting of speech signal helps in reducing the effect of spectral masking for persons with moderate bilateral sensorineural hearing impairment. Speech perception improved by employing filters with interband crossover gain adjusted between 4 and 6 dB below the pass band gain. The relationship between scaling factors for a tone presented to the two ears, so that perceived loudness is that of a monaural presentation, is investigated for design of comb filters with improved perceptually balanced response. Results from the listening tests show that, for perceptual balance, the sum of the two scaling factors should be constant, indicating that the magnitude response of the comb filters should be complementary on a linear scale.

**4aSC36. The organization of bilingual perceptual consonant space: English/Spanish bilingual perception of Malayalam nasal consonants.** Jenna Silver and James Harnsberger (Inst. for Adv. Study of the Commun. Proc., Univ. of Florida, Gainesville, FL 32611)

This study examines the capacity of English/Spanish bilinguals to discriminate between consonants that exist in only one of their respective phonetic inventories. Two non-native nasal consonant contrasts were tested: dental versus alveolar and the palatal versus velar, both found in Malayalam. The dental and palatal nasals appear in Spanish, while the alveolar and velar nasals occur in English. Poorer performance in discrimination was interpreted as indicative of a common nasal category subsuming the Spanish dental and English alveolar nasals; better performance was taken as evidence of the maintenance of separate categories from both languages. Two other tests were administered to aid in the interpretation of the discrimination test scores: forced-choice identification and perceptual similarity ratings. The findings of this research will be used to characterize the perceptual consonant space in terms of continuum between two possible bilingual systems: one that collapses similar categories across languages or one that maintains two distinct phonological systems that can be accessed simultaneously. It is believed that bilinguals will be able to discriminate between these contrasts more consistently than their monolingual peers; however, there is no prediction about performance relative to the monolingual group from Malayalam.

**4aSC37. Agreement and reliability using reference-matching paradigm in perceptual voice quality rating in Chinese and English.** Mandy Ho and Edwin Yiu (Voice Res. Lab., Div. of Speech & Hearing Sci., Univ. of Hong Kong, 5/F Prince Philip Dental Hospital, Hong Kong)

Perceptual voice judgment is commonly used in clinical voice quality evaluation. The use of a referencematching paradigm in perceptual ratings has been shown to improve both agreement and reliability (Yiu *et al.*, in press). This study set out to investigate the agreement and reliability in rating Chinese and English dysphonic stimuli using the reference-matching paradigm. Experiment 1 aimed to synthesize Chinese and English dysphonic stimuli with different breathy and rough severity levels using the HLSyn Speech Synthesizer. Seven representative anchors (references) for each of the rough and breathy series in Chinese and English were chosen by three judges to be used in experiment 2. Acoustic analysis of the anchor series indicated they were of increasing severity. Experiment 2 recruited ten native Chinese and ten native English subjects to rate the quality of Chinese and English dysphonic voice samples using the synthesized anchor as references. Results showed that listeners achieved nearly 90% agreement in rating the Chinese stimuli and 80% agreement in rating the English stimuli regardless of their language background. The study showed that the reference-matching paradigm was a reliable method in rating dysphonic stimuli across listeners with different language backgrounds.

**4aSC38. Learning to perceive non-native speech sounds: The role of test stimulus variability.** McNeel Jantzen and Betty Tuller (Ctr. for Complex Systems and Brain Sci., Florida Atlantic Univ., 777 Glades Rd., Boca Raton, FL 33431)

Natural speech stimuli used in studies of phonological learning usually include several in talkers and phonetic environments because variability aids learning [e.g., Lively, Logan, and Pisoni, *J. Acoust. Soc. Am.* (1993)]. The present study investigated whether nonphonetic variability in the synthetic test set has a similar effect. First, a perceptual mapping procedure was performed using a synthetic continuum that ranged from the Malayalam voiced, unaspirated, dental stop consonant to the American English alveolar [d], with three *F0* contours (low, mid, and high). Subjects identified the stimuli (2AFC) and judged their goodness as exemplars of each category. Subjects then received 15 sessions (one/day) of 2AFC training with feedback using natural stimuli produced by native Malayalam speakers, and performed difference ratings on a subset of pairs from the syn-

thetic stimuli. The perceptual mapping procedure was repeated at 1 and 14 days post-training and results compared with a parallel experiment that included only the midlevel *F0* contour in the synthetic test set. [Work supported by NSF.]

**4aSC39. Influence of the prosody of spoken language on recognition and memory for vocal quality.** Sumi Shigeno (Aoyama Gakuin Univ., 4-4-25 Shibuya, Shibuya-ku, Tokyo, 150-8366 Japan)

This study examined whether recognition and memory for vocal quality of a speaker who speaks either a native language or non-native language should be influenced by the prosody of the language that the speaker utters. Voices of 12 speakers were recorded. They were six Japanese people and six Americans and Britons. All speakers uttered short sentences in their respective native languages (i.e., Japanese for Japanese speakers and English for Americans and Britons) and in a non-native language (i.e., English for Japanese speakers and Japanese for Americans and Britons). Ten Japanese participants rated the vocal quality of speakers in the first session. After 1 week the same experiment was again conducted in the second session. Results showed that the performance of identification of speakers as Japanese or as non-Japanese was comparatively accurate even though the ratings on the speakers' voices were varied as the language spoken by the speakers. Ratings of the voices were compared further between two sessions and little difference was found, irrespective of a 1-week blank. Results suggest that the memory for vocal quality is robust, but that the recognition of vocal quality is dependent on the prosody of the language spoken by speakers.

**4aSC40. Brain activity during auditory processing affected by expectation of speech versus nonspeech.** Yukiko Nota (ATR CIS BAIC, 2-2-2 Hikoridai, Keihanna Sci. City, Kyoto 619-0288, Japan, ynota@atr.jp)

fMRI was used to clarify whether there is any differential brain activity invoked by expectation for speech versus nonspeech sounds. Auditory stimuli were created by acoustically morphing between either sustained vowels or tones, respectively, and a buzz sound. The two sets of interpolation were performed in nine nonlinear steps; the stimuli retained for perceptual experiments were only the three most vowel-like, the three most tone-like, and the three most buzz-like tokens morphed from the vowels. In the "speech expectation" session, subjects were instructed to discriminate the vowel-like and buzz-like stimuli; in the "nonspeech expectation" session, subjects were instructed to discriminate the tone-like and buzz-like stimuli without knowing that the buzz stimuli had been morphed from the vowels. Thus the buzz-like stimuli in both experiments were the same, but the subjects' expectations were different because they were told to expect either speech (vowel-like) or nonspeech (tone-like) stimuli. Comparison of brain activation during processing of the buzz-like stimuli under these two conditions revealed that BA40 and thalamus were more activated in speech expectation, while right BA20 was more activated in nonspeech expectation. These results suggest that subjects' speech/nonspeech expectation for sound stimuli influences brain activity for actual auditory processing. [Work supported by MEXT.]

**4aSC41. Representations involved in short-term versus long-term word learning by preschool children with and without phonological disorders.** Holly Storkel, Jill Hoover, and Junko Maekawa (Dept. of Speech-Lang.-Hearing, Univ. of Kansas, 1000 Sunnyside Ave., 3001 Dole Ctr., Lawrence, KS 66045-7555, hstorkel@ku.edu)

This study explores whether sublexical (i.e., individual sound) and/or lexical (i.e., whole-word) representations contribute to word learning and whether these contributions change across short-term versus long-term learning. Sublexical representations were indexed by phonotactic probability, the likelihood of occurrence of a sound sequence, whereas lexical representations were indexed by neighborhood density, the number of

similar sounding words. Thirty-four preschool children participated in a short-term word learning task that exposed them to nonwords varying in phonotactic probability and neighborhood density and tested learning of these nonwords. Long-term learning was assessed through comprehension and production of real words varying in phonotactic probability and neighborhood density. Results showed that phonotactic probability and neighborhood density equally influenced short-term word learning. In contrast, long-term word learning was affected primarily by neighborhood density. Thus, both sublexical and lexical representations appeared to play a role in short-term learning, but only lexical representations played a primary role in long-term retention. This pattern held for both children with normal phonological development and children with phonological delays. However, the direction of the effect of neighborhood density for short-term word learning varied by group status, suggesting differences in the use of lexical representations during short-term learning. [Work supported by NIH.]

**4aSC42. Changes in formant frequencies associated with postural change in adult male speakers over 50 years old.** Michiko Hashi, Tomoki Nanto, and Natsuki Ohta (Prefectural Univ. of Hiroshima, 1-1 Gakuen-cho, Mihara, Hiroshima, Japan)

It is possible that changes of direction of gravity relative to the vocal tract associated with changes in posture influence acoustic characteristics of speech including vowel formant frequencies. Studies examining such effects had produced mixed results and demonstrated the possibility of substantive interspeaker variability in the effect of postural changes on vowel formant frequencies. Recent work by Takakura *et al.* ["Changes in formant frequencies associated with postural change," paper presented at the Fall meeting of Acoustical Society of Japan (2006)], using young adult male speakers, revealed a small number of speakers demonstrating changes in vowel formant frequencies and suggested effect of age. The present study attempts to examine changes of vowel formant frequencies in upright and supine position among older male speakers. Attempts will be made to eliminate the effect of differences in neck position between the postures through the use of a power-bead-based neck stabilizer. The results will be compared with data from young normal speakers in the previous study and inferences will be made relative to speech production models.

**4aSC43. The effect of viewing angle on the visual contribution to speech intelligibility in noise.** Eugene Brandewie, Douglas Brungart, Nandini Iyer, and Brian Simpson (Air Force Res. Lab., Wright-Patterson AFB, Ohio 45433-7901)

Visual cues are known to assist speech comprehension in noisy environments, but relatively little is known about the impact that viewing angle has on the visual contribution to speech intelligibility. In this experiment, four digital cameras were used to make simultaneous recordings of test phrases from the Modified Rhyme Test at four different viewing angles: 0, 45, 90, and 135 deg. These test phrases were used to measure the effect of viewing angle on the intelligibility of noise-masked speech stimuli that were presented with and without visual cues at seven different signal-to-noise ratios (SNRs). When the face was viewed from the front, the visual cues provided an intelligibility improvement that was equivalent to a 6–10-dB increase in SNR. This visual benefit remained roughly constant for viewing angles up to 90 deg, but it dropped off rapidly (to less than 2 dB) when the viewing angle increased to 135 deg. The results suggest that the visual contribution to speech perception is severely impaired when the observer does not have access to an unobstructed view of the talker's mouth.

**4aSC44. Towards estimation of Japanese intelligibility scores using objective voice quality assessment measures.** Rui Kaga, Kazuhiro Kondo, Kiyoshi Nakagawa, and Masaya Fuzimori (Yamagata Univ., Jounan 4-3-16, Yonezawa, 992-8510, Yamagata, Japan)

We investigated the use of objective quality measures to estimate the intelligibility of Japanese speech. We initially focused on PESQ (perceptual evaluation of speech quality), which is the state-of-the-art objective assessment method, and can estimate the mean opinion scores (MOS) at an extremely high accuracy. Since we can assume that speech quality is correlated with intelligibility, it should be possible to estimate the intelligibility from the estimated opinion scores or some of its derivatives. We tried to estimate the intelligibility of the Japanese Rhyme Test (DRT). The DRT uses minimal word pairs whose initial phone differs by one phonetic feature. We estimated the MOS of the DRT word samples mixed with noise and tried to correlate this with the measured intelligibility. However, the estimated MOS showed no difference between phonetic features. However, the difference in the estimated MOS between the word pairs seems to differ by a phonetic feature for SNR above 10 dB, which suggests that the estimation of intelligibility by a phonetic feature may be possible. We also plan to selectively use the internal data used to calculate the MOS estimate for better intelligibility estimation.

**Session 4aSP****Signal Processing in Acoustics, Underwater Acoustics, and Acoustical Oceanography: Adaptive Signal Processing**

Juan I. Arvelo, Jr., Cochair

*Johns Hopkins Univ., Applied Physics Lab., National Security Technology Dept., 11100 Johns Hopkins Rd., Laurel, MD 20723-6099*

Kensaku Fujii, Cochair

*Univ. of Hyogo, School of Engineering and Graduate School of Engineering, 2-67 Shosha, Himeji, Hyogo 671-2201, Japan***Chair's Introduction—8:00*****Invited Papers*****8:05****4aSP1. Adaptive beamforming for multipath environments.** Henry Cox (Lockheed Martin, IS&S, AC3DI, 4350 North Fairfax Dr., Ste. 470, Arlington, VA 22203)

Coherent multipaths present a significant challenge to most adaptive beamformers because they violate the common assumption of a rank-one plane wave or geometrically focused signal. When the multipath arrivals that characterize shallow water propagation are resolvable by the array's aperture, the mismatch between the assumed and the true signal spatial structure causes signal suppression. If the amplitude and phase relationships among the various multipaths were known, they could, in principle, be included in a matched field beamforming approach. This is usually impractical due to inadequate knowledge of propagation parameters, especially bottom characteristics, and source/receiver motion. A generalization of the standard MVDR approach, called multirank MVDR, assumes only that the signal lies in a subspace of multiple rank rather than the usual rank-one assumption. An example of a subspace is a small fan of beams that cover the potential multipath directions. The signal may be rank one corresponding to fully coherent multipath or higher rank corresponding to incoherent or partially coherent multipath. The multirank approach is applied to the shallow water multipath problem and compared with the related technique of employing multiple linear constraints. Realistic simulations of alternative beamforming approaches for a large horizontal array are presented. [Work supported by ONR.]

**8:25****4aSP2. Robust adaptive algorithm based on nonlinear error cost function for acoustic echo cancellation.** Suehiro Shimauchi, Yoichi Haneda, and Akitoshi Kataoka (NTT Cyber Space Labs., NTT Corp., 3-9-11, Midori-cho, Musashino-shi, Tokyo, 180-8585, Japan, shimauchi.suehiro@lab.ntt.co.jp)

Motivated by recent progress in the blind source separation (BSS) technique, a robust echo cancellation algorithm is investigated, which would inherently identify the echo path even during double-talk by separating the acoustic echo from the local speech. An adaptive filter has been introduced into acoustic echo cancellers to identify the acoustic echo path impulse response and generate the echo replica. However, most adaptive algorithms suffer from instability during double-talk. Although step-size control cooperating with a double-talk detector (DTD) is a promising approach to stop the adaptation temporarily during double-talk, it cannot handle the echo path change during double-talk. To overcome this problem, novel robust algorithms are derived by applying nonlinearity to the cost function of a conventional echo cancellation algorithm such as the normalized least mean squares algorithm (NLMS) or the affine projection algorithm (APA). Using the simulation results, there is a discussion about how the robustness of the derived algorithms depends on the choice of the nonlinear function and the original algorithm.

**8:45****4aSP3. Realistic modeling of adaptive beamformer performance in nonstationary noise.** Bruce K. Newhall (Appl. Phys. Lab., Johns Hopkins Univ., 11100 Johns Hopkins Rd., Laurel, MD 20723)

Most adaptive beamformers (ABFs) operate under the assumption that the noise field is quasistationary. They estimate the present noise field by averaging, assuming stationarity over the estimation time. The adaptive beamformer responds to slow changes in the noise field across multiple estimation intervals. Unfortunately, in many low-frequency underwater sound applications, the shipping noise may change rapidly, due to nearby ship motion. This motion can be significant during the estimation interval and degrade ABF performance. A realistic model has been developed, including two effects of source motion on horizontal towed arrays. Bearing rate produces a differential Doppler shift across the array. Range rate produces an amplitude modulation as the multipath interference pattern shifts along the array. The ABF model begins with a realization of ship locations and motion based on historical shipping density. Each ship generates realistic random noise composed of tonals in a broadband background. The noise is propagated from each ship to each hydrophone by a normal mode model. For each time sample the position of each ship is updated and propagation recalculated. The ability of a variety of ABF algorithms to reduce shipping noise clutter is simulated and compared.

**4aSP4. Multichannel active noise control system without secondary path models using the simultaneous perturbation algorithm.** Yoshinobu Kajikawa and Yasuo Nomura (Faculty of Engineering, Kansai Univ., 3-3-35, Yamate-cho, Suita-shi, Osaka 564-8680, Japan, kaji@joho.densi.kansai-u.ac.jp)

This paper presents a novel multichannel active noise control (ANC) system without secondary path models. This ANC system uses a simultaneous perturbation algorithm as the updating algorithm and has an advantage that secondary path models (estimation of secondary paths) are not required, unlike the MEFX (multiple error filtered-X)-based ANC. This system can consequently control noise stably because there are not modeling errors that cause system instability. The computational complexity is also very small. Experimental results demonstrate that the proposed multi-channel ANC system can operate stably under the environment where the error microphones always move.

### Contributed Papers

9:25

9:55–10:10 Break

**4aSP5. Two-microphone system using linear prediction and noise reconstruction.** Hirofumi Nakano, Kensaku Fujii (Dept. of Comput. Eng., Univ. of Hyogo, 2167 Shosya, Himeji 671-2201, Japan, er06j025@steng.u-hyogo.ac.jp), Tomohiro Amitani, Satoshi Miyata (TOA Corp., Takarazuka Hyogo 665-0043, Japan), and Yoshio Itoh (Tottori Univ., Japan)

This study proposes a new adaptive microphone system that is characterized by a linear prediction circuit inserted previous to a noise reconstruction filter corresponding to an adaptive delay used in conventional systems. This insertion provides various advantages to the adaptive microphone system. One is that steering a null for a noise source becomes possible irrespective of the incidence of speech signal. Another is that the new microphone system works as an omnidirectional microphone for the speech signal. Consequently, setting the microphones at arbitrary intervals is possible. For example, building one microphone in a handset and another into a telephone handset becomes possible, which provides higher noise reduction effect. In practical use, microphone systems must function in reflective surroundings. In this study, such performance of the proposed system is verified first using computer simulations and then using an experimental system put in an ordinary room. This study also presents experimental results verifying that the proposed system can successfully reduce noise incident from the same direction as a speech signal, as well as crowd noise recorded in an airport.

10:10

**4aSP7. Adaptive matched field processing enhancements to forward sector beamforming.** Jeffrey A. Dunne (Appl. Phys. Lab., Johns Hopkins Univ., 11100 Johns Hopkins Rd., Laurel, MD 20723)

A study was undertaken to examine the potential benefit of adaptive matched field processing (AMFP) to the forward sector capability of single-line, twin-line, and volumetric arrays. Comparisons are made with conventional MFP (CMFP) and adaptive and conventional plane-wave beamforming (APWB and CPWB) in order to assess the degree of ownership noise reduction obtainable and any corresponding improvement to the signal-to-noise ratio (SNR). A minimum variance distortionless response beamformer using dominant mode rejection was implemented, applied to both uniform and distorted array shapes. Significant improvement over CMFP and CPWB in tracking and SNR was seen for modeled data in both cases, with the distorted array showing, not surprisingly, better left-right rejection capability. [Work was undertaken with support from the Defense Advanced Research Projects Agency (DARPA) Advanced Technology Office (ATO).]

10:25

9:40

**4aSP6. Adaptive beamformer trade-off study of an expendable array for biologic vocalizations.** Juan Arvelo (Appl. Phys. Lab., Johns Hopkins Univ., 11100 Johns Hopkins Rd., Laurel, MD 20723-6099)

Adaptive beamformers exploit the ambient noise anisotropy to increase the array gain against this background, enhance target detection, and increase the resolution of the beam response. A prototype expendable array was developed and tested for high-frequency passive detection and localization of marine mammals. This latest array consists of vertical poly(vinylidene fluoride) (PVDF) wire elements arranged in a horizontal  $6 \times 6$  grid with the corner elements removed for a total of 32 lines. The length of the wires forms a vertical beam response that allows exploitation of the ambient noise directionality in elevation while the horizontal aperture provides full coverage in azimuth. The performance and computational demand of selected adaptive and conventional beamformers are compared in a trade-off study to determine their possible use in this expendable system. This trade-off study accounts for the demand of computational resources in addition to the predicted system performance as adjuncts to ocean observatories. [This effort is partly supported by JHU/APL and the Office of Naval Research (ONR).]

**4aSP8. Vector sensor array sensitivity and mismatch: Generalization of the Gilbert-Morgan formula.** Andrew J. Poulsen and Arthur B. Baggeroer (MIT, Cambridge, MA 02139, poulsen@mit.edu)

The practical implementation of any sensing platform is susceptible to imperfections in system components. This mismatch or difference between the assumed and actual sensor configuration can significantly impact system performance. This paper addresses the sensitivity of an acoustic vector sensor array to system mismatch by generalizing the approach used by Gilbert and Morgan for an array with scalar, omnidirectional elements [E.N. Gilbert and S.P. Morgan, *Bell Syst. Tech. J.* **34**, (1955)]. As such, the sensor orientation is not an issue because it does not affect performance for an array of omnidirectional sensors. Since vector sensors measure both the scalar acoustic pressure and acoustic particle velocity or acceleration, the sensor orientation must also be measured to place the vector measurement in a global reference frame. Here, theoretical expressions for the mean and variance of the vector sensor array spatial response are derived using a Gaussian perturbation model. Such analysis leads to insight into theoretical limits of both conventional and adaptive processing in the presence of system imperfections. Comparisons of theoretical results and simulations are excellent. One noteworthy result is the variance is now a function of the steering angle. [Work supported by the PLUSNet program of the Office of Naval Research.]

10:40

**4aSP9. Adaptive filtering using harmonic structure of voiced speech for reducing nonstationary known noise.** Kenko Ota, Masuzo Yanagida (Doshisha Univ., 1-3, Tarata-Miyakodani, Kyotanabe, Kyoto, 610-0321, Japan, etf1704@mail4.doshisha.ac.jp), and Tatsuya Yamazaki (NICT, 619-0289, Kyoto, Japan, yamazaki@nict.go.jp)

Proposed is an effective method for reducing nonstationary known noise. The objective of this research is to develop a scheme of preprocessing for speech recognition that keeps the same speech recognition rate even in a worse acoustic environment and to realize a TV control system using speech recognition. The basic idea of the proposed method is to estimate the frequency spectrum of noise including sounds from the TV itself and to remove noise components from the frequency spectrum of the received signal. A transfer function from the TV set to the microphone is calculated in an iterative manner estimating the noise signal at the microphone. Traditional ANC techniques do not directly use spectral features such as harmonic structure or fundamental frequency, for example. As the proposed method uses harmonic structure of vocalic segments in command speech, it is expected to have an advantage compared with the traditional adaptive methods. Results of evaluation using speech recognition show that the proposed method significantly improves speech recognition rate compared with conventional spectral subtraction. [Work supported by Knowledge Cluster Project, MEXT, and by Academic Frontier Project Doshisha University.]

10:55

**4aSP10. Robust matched-field processor in the presence of geoacoustic inversion uncertainty.** Chen-Fen Huang, Peter Gerstoft, and William S. Hodgkiss (Marine Physical Lab., Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92037-0238, chenfen@ucsd.edu)

This presentation examines the performance of a matched-field processor incorporating geoacoustic inversion uncertainty. Uncertainty of geoacoustic parameters is described via a joint posterior probability distribution (PPD) of the estimated environmental parameters, which is found by formulating and solving the geoacoustic inversion problem in a Bayesian framework. The geoacoustic inversion uncertainty is mapped into uncertainty in the acoustic pressure field. The resulting acoustic field uncertainty is incorporated in the matched-field processor using the minimum variance beamformer with environmental perturbation constraints (MV-EPC). The constraints are estimated using the ensemble of acoustic pressure fields derived from the PPD of the estimated environmental parameters. Using a data set from the ASIAEX 2001 East China Sea experiment, tracking performance of the MV-EPC beamformer is compared with the Bartlett beamformer using the best-fit model.

11:10

**4aSP11. A study on combining acoustic echo cancelers with impulse response shortening.** Stefan Goetze, Karl-Dirk Kammeyer (Dept. of Commun. Univ. of Bremen, Eng., D-28334 Bremen, Germany), Markus Kallinger, and Alfred Mertins (Carl von Ossietzky-Univ., Oldenburg, D-26111 Oldenburg, Germany)

In hands-free video conferencing systems acoustic echo cancelers (AECs) have to face the problem of very high-order impulse responses (IRs), which have to be compensated. Time-domain algorithms for adaptation often suffer from slow convergence (as the NLMS algorithm, e.g.) or high computational complexity (e.g., the RLS). On the other hand frequency-domain algorithms introduce undesired delays [S. Haykin, *Filter Theory*, 2002]. For high-quality hands-free systems IR shortening concepts and IR shaping concepts developed for listening room compensation

(LRC) [M.Kallinger and A. Mertins, in Proc. Asilomar, 2005] can be applied to increase speech intelligibility for the near-end speaker. The aim of this study is the synergetic combination of LRC concepts with acoustic echo cancellation. For this scenario two different forms of concatenating the subsystems are possible. Either the AEC filter follows the LRC or vice versa. In the first case the equalization filter reduces the length of the effective IR seen by the AEC filter. Thus, shorter AEC filters can be used which results in faster convergence. However, an estimation algorithm for the room IR is necessary for the LRC subsystem. In the second case the AEC delivers an estimate of the room IR which can be used as an input for the LRC filter. Experimental results confirm the superiority of the new combined approach.

11:25

**4aSP12. Krylov and predictive sequential least-squares methods for dimensionality reduction in adaptive signal processing and system identification.** James Preisig and Weichang Li (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Rapid time variation of the environment, a large number of parameters which need to be adjusted, and the presence of a reduced subset of the parameters that are relevant at any point in time create significant challenges for adaptive signal-processing algorithms in underwater acoustic applications. In applications such as underwater acoustic communications, the environment is represented by the "taps" of the time-varying impulse response. The instability of estimation algorithms or inability to track rapid channel fluctuations are among the problems that are encountered. An approach to addressing these challenges is to dynamically select a subspace in which the adjustment of taps takes place. Here, two algorithms for doing this are presented. The first is based upon using subspace basis vectors, which form a Krylov subspace with respect to the channel input correlation matrix and the channel input/output correlation vector. This method does not use a prediction residual error to select the basis vectors. A second algorithm is a new variant of the matching pursuit algorithm. In this case, "important" taps of the channel impulse response are selected to minimize a forward prediction residual error. The properties and performance of these two algorithms are presented and compared using simulation and field data.

11:40

**4aSP13. Expectation maximization joint channel impulse response and dynamic parameter estimation and its impact on adaptive equalization.** Weichang Li and James C. Preisig (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanograph. Inst., Woods Hole, MA 02543)

Joint estimation of channel impulse response and its dynamic parameters using the expectation maximization (EM) algorithm and its MAP variant is derived for broadband shallow-water acoustic communication channels. Based on state-space channel modeling, the EM algorithms estimate the channel dynamic parameters from the sequence of channel impulse response estimates. The estimated parameters are then used in the Kalman smoother, which estimates the channel impulse response. The stability of the algorithm is shown to be related to an extended persistent excitation (EPE) condition, which requires that both the symbol sequence and the channel estimates be persistently exciting. Modified algorithms are proposed for broadband multipath channels to avoid the issue of insufficient excitation. Efficient suboptimal algorithms are also derived from the EM algorithms that alternatively estimate the parameter and the channel impulse response while allowing slow parameter variations. The performance of these channel estimation algorithms as well as their impact on the subsequent equalizer are demonstrated through experimental data analysis. [Work supported by ONR Ocean Acoustics.]

## Session 4aUWa

## Underwater Acoustics: Sonar Performance

Lisa M. Zurk, Cochair

*Portland State Univ., Electrical and Computer Engineering Dept., 1900 S. W. Fourth Ave., Portland, OR 97207*

Hiroshi Ochi, Cochair

*Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-1-5 Natsushima-cho, Yokosuka, Kanagawa 237-0061, Japan*

Chair's Introduction—7:55

## Contributed Papers

8:00

**4aUWa1. Automatic detection performance comparisons of three different fluctuation-based signal processors.** Ronald A. Wagstaff (Natl. Ctr. for Physical Acoust., 1 Coliseum Dr., Univ., MS 38677)

The three most common fluctuation-based signal processor (FBP) algorithms achieve gain by exploiting either the reciprocal of the spectral amplitude, the log-differential amplitude, or aligned-phase angles. Two important features of these processors, for the underwater acoustics community, is their ability to detect and automatically identify signals which originated from submerged sources, and to provide unusually large signal-to-noise ratio gains. Similar benefits are of interest to the atmosphere acoustic community. An example is the automatic detection and identification/classification of hostile airborne and ground vehicles by unattended ground sensors (UGS). The three different generic types of FBP algorithms will be defined. The manner in which each exploits fluctuations to achieve gain will be explained. Corresponding performances will be compared using both underwater and atmosphere acoustic data. The ocean acoustic results came from towed array beamformed spectral data and will include spectral plots and grams. Corresponding single sensor spectral results will also be presented for atmosphere acoustic vehicle data. [Work supported by ARDEC and SMDC.]

8:15

**4aUWa2. Evolution of modern fluctuation-based processing.** Ronald Wagstaff (Natl. Ctr. for Physical Acoust., 1 Coliseum Dr., Univ., MS 38677)

Modern fluctuation-based processors (FBPs) are relatively new on the signal processing scene. They started in the mid-1980s with the realization that averaging the reciprocal acoustic powers and inverting the result back, i.e., the harmonic mean, could yield 6- to 8-dB signal-to-noise ratio gain over the corresponding average power. Because of its significant noise attenuation, this processor was designated WISPR, i.e., reduces noise to a whisper. WISPR had a unique, potentially more valuable, capability. Based on the decibel difference, or ratio, between the average power and WISPR, it could be determined whether the received signals were from ships, or from sources of sound that were submerged. After much time and experience with WISPR at sea, acquiring and processing towed array ocean acoustic data, and continuing data processing in the laboratory, the phenomena that were responsible for WISPRs performance, acoustic fluctuations generated near the sea surface, became better understood and WISPRs credibility increased. This led to the development of many other FBPs with similar capabilities, but with significantly enhanced performances. A brief account of post-WISPR development will be presented, including a description of the exploitable parameters, how they are used, and the range of gains that they achieve.

8:30

**4aUWa3. A comprehensive unbiased third party evaluation of a signal processor for detecting submerged sources among clutter signals and noise.** Ronald Wagstaff (Natl. Ctr. for Physical Acoust., 1 Coliseum Dr., Univ., MS 38677)

The Wagstaff's integration silencing processor, WISPR', was developed to detect and identify signals in the ocean from sources that are submerged well below the sea surface. WISPR is the type of signal processor that exploits the reciprocal of the spectral power amplitude, rather than the amplitude as the average power processor does. Processing the reciprocal of the power represented a significant departure in the prevailing signal processing philosophy that governed most conventional signal processing algorithms that were in use when WISPR first appeared on the scene several years ago. Furthermore, WISPR's claimed submerged-source detection capability made it an attractive candidate for some high interest signal processing applications. Accordingly, one influential national organization considered its potential use in their mission and decided to commission a credible third party laboratory to conduct an unbiased evaluation of the WISPR processor. The emphasis was to be on its performance for automatic unalerted detection of signals from submerged sources. The techniques and evaluation methods used to test the WISPR processor will be described. The results of the evaluation will be presented, and the influence of those results on the development of other, more advanced, fluctuation-based processors will be discussed.

8:45

**4aUWa4. The estimated ocean detector: Predicted performance for continuous time signals in a random/uncertain ocean.** Jeffrey A. Ballard, R. Lee Culver, Leon H. Sibul (Appl. Res. Lab. and Grad. Program in Acoust., Penn State Univ. P.O. Box 30, State College, PA 16804), Colin W. Jemmott, and H. John Camin (Penn State Univ., State College, PA 16804)

This paper addresses implementation of the maximum likelihood (ML) detector for passive SONAR detection of continuous time stochastic signals that have propagated through a random or uncertain ocean. We have shown previously that Monte Carlo simulation and the maximum entropy method can make use of knowledge of environmental variability to construct signal and noise parameter probability density functions (pdf's) belonging to the exponential class. For these cases, the ML detector has an estimator-correlator and noise-canceller implementation. The estimator-correlator detector computes the conditional mean estimate of the signal conditioned on the received data and correlates it with a function of the received data, hence the name estimated ocean detector (EOD). Here we derive the detector structure for continuous time stochastic signals and Gaussian noise and present receiver operating characteristic (ROC) curves for the detector as a function of the signal-to-noise ratio. [Work supported by ONR Undersea Signal Processing Code 321US.]

9:00

**4aUWa5. Echo detection enhancement using multiple guide sources in shallow water.** David C. Calvo, Charles F. Gaumont, David M. Fromm, and Richard Menis (Naval Res. Lab., Acoust. Div. Code 7145, 4555 Overlook Ave. SW, Washington, DC 20375)

The use of a guide source has been proposed as a way of compensating for multipath by forming a spatial-temporal cross correlation of the received target and guide source signals across a vertical array in the frequency domain [Siderius *et al.*, *J. Acoust. Soc. Am.* **102**, 3439–3449]. This processing has the effect of creating a virtual receiver at the guide source position. In general, the performance of a virtual receiver degrades if the spatial integration is not carried out over the span of the array with significant signal. In our study, we have pursued an alternative approach of using guide sources which does not require this integration in general. The guide source signal is simply used as a matched filter. Although this does not correspond to a virtual receiver, it is useful as a means of improving active or passive detection of signals in noise. In general, the signal gain using this alternative technique is dependent on the guide source position. To compensate for this, we construct a separable-kernel-receiver filter bank using multiple randomly positioned guide source signals. Improvement of ROC curves in both passive and active scenarios is obtained using experimental and simulated data. [Work sponsored by ONR.]

9:15

**4aUWa6. Incorporating environmental variability into received signal statistics.** H. John Camin, R. Lee Culver, Leon H. Sibul (Appl. Res. Lab. and Grad. Program in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804), Jeffrey A. Ballard, and Colin W. Jemmott (Penn State Univ., State College, PA 16804)

We have developed a Monte Carlo-based method for estimating the variability of acoustic signal parameters caused by uncertain ocean environments. The method begins with a physics-based model for the environmental properties and uses the maximum entropy (MaxEnt) method to construct probability density functions (pdf's) describing the measured deviations from the model mean. Random realizations of environmental variability, with proper depth correlation, are constructed from the pdf's and added to the mean model parameters. A parabolic equation code (RAM) is used to propagate acoustic energy through each realization of the environment. Fourier synthesis is used to recreate the arrival structure. The method is demonstrated using measurements from the Strait of Gibraltar, which is a particularly complicated region dominated by strong tidal fluctuations and internal waves. During 1996, an international group carried out the Strait of Gibraltar Acoustic Monitoring Experiment (SGAME), in which detailed environmental and 250-Hz acoustic data were collected.

Here, pdf's of the received signal level are compared with results of the Monte Carlo method to demonstrate performance. [Gibraltar data and SVP model provided by Chris Tiemann (ARL:UT) and Peter Worcester (SIO). Work supported by ONR Undersea Signal Processing.]

9:30

**4aUWa7. Motion compensation of multiple sources.** Joung-Soo Park (Agency for Defense Development, P.O. Box 18, Jinhae, Kyung-Nam, 645-600, Korea), Jae-Soo Kim (Korea Maritime Univ., Young-Do, Busan, Korea), and Young-Gyu Kim (Agency for Defense Development, Jinhae, Kyung-Nam, 645-600, Korea)

Matched field processing has a advantage of detection of multiple targets. But, if a strong interferer is moving fast near a quiet target, detection of the target is difficult due to the motion effect of the interferer. The motion of the interferer introduces energy spreading and results in poorer detection. A waveguide-invariant-based motion compensation algorithm was proposed to mitigate the motion effect of a dominant signal component, which is estimated by eigenvalue method. The eigenvalue method is good for a strong interferer, but not good for multiple targets. In this presentation, we will propose a steered beam processing method to mitigate the motion effect of multiple targets. We will verify the proposed method with numerical simulations and SwellEx96 data processing.

9:45

**4aUWa8. Predicting sonar performance using observations of mesoscale eddies.** Harry DeFerrari (Div. of Appl. Marine Phys., RSMAS, Univ. of Miami, 4600 Rickenbacker Cswy, Miami, FL 33149)

A predictive relationship has been observed between the location of offshore mesoscale eddies and the performance of active and passive sonar on the shallow water shelf area inside of the eddy. The passage of an eddy produces a prograde front that modifies acoustic propagation by two mechanisms. First, the density gradient serves as a conduit for offshore internal waves to propagate onto the shelf. A long-lived front can result in order of magnitude increases in potential energy of the internal wave field and corresponding increases in sound speed variability. Second, the circulation of the eddy produces a unique sound speed profile that is strongly downward refracting but has a nearly iso-velocity layer near the bottom owing to turbulent mixing. The shape of the profile closely approximates a hyperbolic cosine. Such a profile has mode group velocities that are equal for all refracted modes, thus producing strong focusing and a caustic at the depth of the source at all ranges. The experimental observations are confirmed with oceanographic and acoustic propagation models and, in turn, the models predict FOM fluctuations of as much as 15 dB for passive sonar and 24 dB for active sonar, depending on location of the eddy.

## Session 4aUWb

## Underwater Acoustics: Session in Honor of Leonid Brekhovskikh I

William A. Kuperman, Cochair

*Scripps Inst. of Oceanography, Univ. of California, San Diego, Marine Physical Lab., MC0238, San Diego, La Jolla, CA 92093-0238*

Oleg A. Godin, Cochair

*NOAA, Earth System Research Lab., 325 Broadway, Boulder, CO 80305-3328*

Chair's Introduction—10:15

*Invited Papers*

10:20

**4aUWb1. Phenomenon of Leonid Maximovich Brekhovskikh as a man and a scientist.** Nikolay Dubrovskiy (Andreyev Acoust. Inst., Shvernika Ul 4, Moscow 117036, Russia)

Leonid Maximovich Brekhovskikh made phenomenal contributions in acoustics: discovery of the underwater sound channel, development of the fundamental theory of wave propagation in layered media, and working out a tangent plane approximation in the wave scattering theory. Brekhovskikh contributed greatly to the organization of research teams and the dissemination of information on acoustics and oceanography through his popular books and lecturing. He also made a major breakthrough as a public figure and a statesman. He became the first Director of the Acoustics Institute at the age of 36. He served as a secretary of the Russian Academy of Sciences Branch involved in oceanography, geography, and atmospheric physics research. Brekhovskikh's achievements in science and science leadership were marked by multiple highest USSR awards and many international awards. He became an Honorary Fellow of the Acoustical Society of America and the Russian Acoustical Society. He received the Lord Raleigh Medal for the discovery that preserved its urgency for 30 years. Brekhovskikh's phenomenon is regarded here from the viewpoint of his personality as well as specific circumstances of his family and social life.

10:40

**4aUWb2. Some aspects of Leonid Brekhovskikh's influence on oceanographic acoustics.** W. A. Kuperman and W. Munk (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0238)

Waveguide physics describes the basic features of long-range sound propagation in the ocean. Over the last half century the theory has progressed from describing ideal waveguides to more complicated layered structures to range-dependent structures to time-varying, range-dependent structures. The theme of Brekhovskikh's pioneering work was the development of robust formulations that permitted understanding basic ocean acoustics while also laying the foundation to progress to the next levels of realistic complexity. Early on, he realized that acoustic data were not consistent with known oceanography. His seminal oceanographic experiments established the pervasive presence of mesoscale phenomena, which to this day are still not fully incorporated into rigorous formulations of the forward and inverse acoustics problems. We discuss only a very small part of his work and its subsequent influence.

11:00

**4aUWb3. Underwater sound propagation: 49 years with L. M. Brekhovskikh's *Waves in Layered Media*.** Oleg A. Godin (CIRES, Univ. of Colorado and NOAA, Earth System Res. Lab., 325 Broadway, Boulder, CO 80305, oleg.godin@noaa.gov)

In his first 10 years of research on wave propagation in layered media, L. M. Brekhovskikh created a theory that remains a basis for physical understanding and mathematical modeling of underwater sound propagation. Summarized in his celebrated book *Waves in Layered Media*, first published in 1957, the theory includes spectral (quasi-plane wave) representations of wave fields, normal mode theory for open waveguides, extensions of the ray and WKBJ methods, and a clear treatment of diffraction phenomena attendant to caustics, lateral waves, and reflection of wave beams and pulses. The book also charted the ways forward that have been and are followed by numerous researchers around the globe. Some of the resulting progress was documented in subsequent editions of *Waves in Layered Media* and in later books L. M. Brekhovskikh coauthored with his students. This paper will discuss diverse, groundbreaking contributions L. M. Brekhovskikh made to the wave propagation theory from the perspective offered by recent developments in underwater acoustics.

11:20

**4aUWb4. L. M. Brekhovskikh's studies on nonlinear wave interaction and atmospheric sound.** Konstantin Naugolnykh (Univ. of Colorado, NOAA, ESRL/Zeltaech LLD, Boulder, CO)

Nonlinear interaction of waves in a compressible fluid is an underlying factor in many geophysical effects, and L. M. Brekhovskikh made essential contributions to investigation of these phenomena. In particular, he suggested the mechanism of the infrasound generation by stormy areas in the ocean based on the nonlinear interaction of the counter-propagating sea-surface gravity

waves. The estimates of the order of magnitude of sound intensities were made indicating that the main part of the infrasound generated by the surface waves is absorbed in the upper layers of the atmosphere, resulting in the heating of these layers. The other part of the sound energy can be trapped by the atmospheric acoustic waveguide and then returned to earth at distances of hundreds of kilometers, producing the voice of the sea.

11:40

**4aUWb5. Tangent-plane approximation by L. M. Brekhovskikh and connected methods in the theory of wave scattering from rough surfaces.** Alexander G. Voronovich (NOAA, Earth System Res. Lab., Physical Sci. Div., 325 Broadway, Boulder, CO 80305, alexander.voronovich@noaa.gov)

Starting from pioneering work by Rayleigh in 1907, scattering of waves from rough surfaces was restricted by the case of small Rayleigh parameter. In this case perturbation analysis describing the process of Bragg scattering applies. Apparently, smallness of the roughness is too restrictive for many applications. In 1952 L. M. Brekhovskikh suggested a tangent-plane approximation (TPA). For ideal boundary conditions it represents the first iteration of the appropriate boundary integral equation. However, for more complex situations (e.g., dielectric or solid-fluid interfaces) appropriate boundary integral equations are rather complicated and, even worse, they cannot be readily iterated. The TPA allows bypassing this step providing the answer in closed form for arbitrary boundary conditions and for scalar or vector waves in terms of the local reflection coefficient. Unfortunately, the TPA does not correctly describe the Bragg scattering. However, later it was realized that the TPA allows simple generalization, which treats both low- and high-frequency limits within single theoretical scheme. This is achieved by considering the local reflection coefficient as an operator rather than a factor. New methods going beyond the two classical ones with much wider regions of validity were developed based on this idea. Some of them will be reviewed in this talk.

FRIDAY AFTERNOON, 1 DECEMBER 2006

LANAI ROOM, 1:00 TO 4:45 P.M.

### Session 4pAA

## Architectural Acoustics: Measurement of Room Acoustics II

Boaz Rafaely, Cochair

*Ben Gurion Univ., Electrical and Computer Engineering Dept., 84105 Beer Sheva, Israel*

Hideo Miyazaki, Cochair

*Yamaha Corp., Ctr. for Advanced Sound Technologies, 203 Matsunokijima, Iwata, Shizuoka 438-0192, Japan*

### Contributed Papers

1:00

**4pAA1. Impulse response measurements based on music and speech signals.** Wolfgang Ahnert, Stefan Feistel, Alexandru Miron, and Enno Finder (Ahnert Feistel Media Group, Berlin, Germany)

All known software based measurement systems, including TEF, MLSSA, SMAART, and EASERA, derive results using predetermined excitation signals like Sweep, MLS, or Noise. This work extends the range of excitations to natural signals like speech and music. In this context selected parameters like frequency range, dynamic range, and fluctuation of the signal and the signal duration are investigated in order to reach conclusions about the conditions required to obtain results comparable with standard excitation signals. Also the limitations of the standard stimuli and the proposed natural stimuli are discussed.

1:15

**4pAA2. Assessment of reverberation time in halls through analysis of running music.** David Conant (McKay Conant Brook Inc., 5655 Lindero Canyon Rd., Ste. 325, Westlake Village, CA 91362, dconant@mcbinco.com)

The source signal to excite a room's reverberant field sufficient for detailed measurement of reverberation time (RT60) and other measures has been the subject of considerable investigation over several decades. It is generally acknowledged that the best sources are (depending on the researcher) swept tones, MLS, MLS variations, stopped noise, cannon shots, etc. All can be characterized as highly audience unfriendly. In the

interest of obtaining useful approximations of measured midfrequency RT60 in the presence of live audiences, this paper discusses several approaches that may be fruitful while being entirely unobtrusive to the concert experience.

1:30

**4pAA3. Comparison of measurement techniques for speech intelligibility.** Bruce C. Olson (Olson Sound Design, 8717 Humboldt Ave N, Brooklyn Park, MN 55444, bco@olsonsound.com)

A comparison of measurement techniques for speech intelligibility between two recently released measurement systems is made. EASERA (Electronic and Acoustic System Evaluation and Response Analysis) uses a standard PC and an EASERA Gateway interface attached via Firewire. The software postprocesses a variety of stimuli in order to derive the impulse response for the room under test. This impulse response is then further processed and the results are presented to the user in both graphical and textual presentations. The Ivie Technologies IE-35 is based on a Pocket PC system and uses an external modulated noise source as stimulus to produce an intelligibility score as a single number or average of a series of measurements. This paper will explore a variety of measurements made in the same locations in a room by both systems. Results will also be shown for a variety of other acoustic measures that quantify the acoustical parameters of the room.

**4pAA4. Under-balcony acoustics in concert halls: Single source versus an array of multiple sources.** Youngmin Kwon and Gary W. Siebin (Architecture Technol. Res. Ctr., Univ. of Florida, 134 ARCH, P.O. Box 115702, Gainesville, FL 32611, ymkwon@ufl.edu)

The conventional measurement protocol using a single omnidirectional sound source may have limits or uncertainty in objective acoustical analysis of a performance hall. This study conducted monaural and binaural impulse response measurements with an array of 16 directional loudspeakers for quantitative acoustical assessment of specifically under-balcony area in a concert hall. The measurements were executed in a real performance hall. The measured time- and frequency-domain responses as well as the results of room acoustical parameters including binaural parameters were compared to the ones measured with a single omnidirectional source. The results were also compared to the ones taken at the main orchestra seating area. The time-domain responses showed a clear distinction particularly in early responses between single source and multiple sources. On the other hand, the magnitude of frequency response showed significantly lower at frequencies above 1 kHz than the one measured at the main area. The results of a binaural parameter, IACC, were found to be marginal between single source and multiple sources but critically different between under-balcony area and main area. Variations were also observed in the results of other room acoustical parameters when compared either between single source and multiple sources or between under-balcony area and main area.

2:00

**4pAA5. Alternative metrics for the directivity of acoustic sources.** Timothy W. Leishman (Acoust. Res. Group, Dept. of Phys. and Astron., Brigham Young Univ., Provo, UT 84602)

While the directivity of an acoustic source at a given frequency is thoroughly characterized by a directivity function over the angular coordinates, it may also be characterized to a lesser degree by a single-number directivity factor. The directivity index (i.e., the logarithmic version of the directivity factor) is a related figure of merit. Recent efforts to quantify the directivity of sources for architectural acoustics measurements have led to several alternatives to these values. One example is the area-weighted spatial standard deviation of radiated levels over a free-field measurement sphere. This paper presents and compares this and other directivity metrics for several types of sources, and discusses their benefits.

2:15

**4pAA6. Room volume estimation from diffuse field theory.** Martin Kuster and Maarten van Walstijn (Sonic Arts Res. Ctr., Queen's Univ. Belfast, BT7 1NN Belfast, Northern Ireland, m.kuster@qub.ac.uk)

Among the parameters relevant in room acoustics, the room volume is one of the most important. The general course in room acoustics research is to use the room volume in the prediction of room acoustic parameters such as reverberation time or total relative sound pressure level. Contrary to this, it has been investigated to what extent the room volume can be retrieved from a measured room impulse response. The approach followed is based on room acoustic diffuse field theory and requires correctly measured room impulse responses with the initial time delay corresponding to the source to receiver distance. A total of ten rooms of varying size and acoustic characteristics have been included. The results in three rooms were unreliable, which was explained by the particular acoustic characteristics. In the remaining rooms the results were numerically useful and consistent between different positions within the same room (relative standard deviation around 20%). The influence of source and receiver directivity is also considered.

**4pAA7. In situ measurements for evaluating the scattering surfaces in a concert hall.** Jin Yong Jeon and Shin-ichi Sato (School of Architectural Eng., Hanyang Univ., Seoul 133-791, Korea, jyjeon@hanyang.ac.kr)

Sound diffusion by a wall structure is one of the main concerns with respect to the sound quality of concert halls. There is a need to develop measurement and evaluation methods for determining the performance of scattering wall surfaces not only in a laboratory but also in actual halls. In this study, the acoustical measurements were conducted in a concert hall which has diffusers with ceramic cubic tiles on the side walls of the stage and the audience area. Binaural impulse responses were measured at all of the seats under two conditions, that is, with and without diffusers. The area which was affected by the diffusive wall was determined and quantified. The condition without diffusers was produced by covering them with the movable reflectors. From the binaural impulse responses, the temporal diffusion [H. Kuttruff, *Room Acoustics*, (Elsevier Science, London, 1991)], which is calculated from the autocorrelation of the impulse response, and other acoustical parameters were analyzed. From the relationship between the scattering coefficient and the acoustical parameters, sound scattering index for real halls, which represents the degree of the diffusion of a hall, was proposed.

2:45

**4pAA8. Further investigations on acoustically coupled spaces using scale-model technique.** Zuhre Su, Ning Xiang (Grad. Program in Architectural Acoust., School of Architecture, Rensselaer Polytechnic Inst., Troy, NY 12180), and Jason E. Summers (Naval Res. Lab., Washington, DC 20024)

Recently, architectural acousticians have been increasingly interested in halls that incorporate coupled-volume systems because of their potential for creating nonexponential sound energy decay. Effects of coupling-aperture configuration and source and receiver locations on energy decay are essential aspects of acoustically coupled spaces that have not yet been extensively investigated. In order to further understand these effects on sound fields in coupled rooms, a systematic experimental study is carried out. An acoustic scale model technique is used in collecting room impulse responses of a two-room coupled system for varying aperture configurations and surface-scattering conditions. Baseline behavior is established by varying aperture area for a fixed aperture shape and analyzing relevant energy-decay parameters at different locations. Effects of aperture shape and number are systematically investigated by varying these parameters while holding coupling area fixed. Similarly, effects of receiver location are systematically investigated by varying the distance of the receiver from the coupling aperture for a fixed aperture configuration. Schroeder decay-function decompositions by Bayesian analysis reveal sensitivities to receiver location and aperture configuration across different frequency bands.

3:00–3:15 Break

3:15

**4pAA9. Virtual microphone control: A comparison of measured to created impulse responses of various microphone techniques.** Daniel Valente and Jonas Braasch (Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180, danvprod@yahoo.com)

A method of rendering sound sources in 3-D space has been developed using virtual microphone control (ViMiC) [J. Acoust. Soc. Am. **117**, 2391]. This method has been used to create a flexible architecture for the creation and rendering of a virtual auditory environment based on microphone techniques. One of the advantages of ViMiC is the ability to simulate coincident, near-coincident, and spaced microphone recording techniques. This allows the user active spatial control over the recorded environment and the ability to shape the final rendering based on his or her specific auditory needs. In order to determine the accuracy of simulating the virtual microphone techniques, measurements of several acoustic

spaces in Troy, NY will be compared to the measurements of generated impulse responses of the same modeled spaces within the ViMiC environment. The data from the measured impulse responses will be used to adapt the ViMiC system in order to create a more realistic auditory rendering. Moreover, the ViMiC system can be improved for use as an educational tool for teaching recording engineers to hear the subtle differences between various microphone techniques.

3:30

**4pAA10. The estimation of the room acoustic characteristic using the acoustic intensity method.** Yong Tang, Hideo Shibayama, and Takumi Yosida (Dept. of Commun. Eng., Shibaura Inst. of Technol., 3-7-5 Toyosu Koutou-ku, Tokyo, 135-8548 Japan, m603101@sic.shibaura-it.ac.jp)

When a sound radiates in rooms, a lot of reflection sounds are generated. From estimation of the direction where the room reflection sound comes from, we can understand the diffusion situation in the room acoustic field. By using the acoustic intensity method, we can measure the strength and the direction of the sound. In this paper, we estimate the direction of the reflection sound in the time-space by the acoustic intensity method and show the acoustic characteristic of the room.

3:45

**4pAA11. Binaural simulation in an enclosure using the phased beam tracing.** Cheol-Ho Jeong and Jeong-Guon Ih (NOVIC, Dept. of Mech. Eng., KAIST, Sci. Town, Daejeon 305-701, Korea, chjeong@kaist.ac.kr)

Binaural simulation in an enclosed space is important in the subjective evaluation of the enclosure acoustics in the design or refinement stage. A time domain scheme using the geometrical acoustics technique has been usually used in the binaural processing. However, one can calculate a pressure impulse response by using the phased beam-tracing method, which incorporates the phase information in the beam tracing process. Such phased method employs reflection coefficient and wave number, whereas the conventional method uses absorption coefficient and air attenuation factor. Impulse response can be obtained by the inverse Fourier transformation of the frequency domain result. This feature facilitates the binaural simulation because the convolution with the HRTF can be accomplished by a simple multiplication in frequency domain. Convolutions were conducted for all reflections one by one, and the convolved transfer functions were summed into one transfer function. Consequently binaural room impulse responses at receivers' ear positions can be simulated. The measured binaural room impulse responses in the conference room were compared with the predicted results for octave bands of 125 Hz to 4 kHz. A good agreement with measurement was found, especially in the early part of impulse responses. [Work supported by BK21.]

4:00

**4pAA12. Visualization methods of direct and early reflection sounds in small enclosures.** Chiaki Koga, Akira Omoto (Omoto Lab., Dept. of Acoust. Design, Faculty of Design, Kyushu Univ., Shiobaru 4-9-1, Minami, Fukuoka 815-8540, Japan), Atsuro Ikeda, Masataka Nakahara (SONA Corp., Nakno-ku, Tokyo, 164-0013, Japan), Natsu Tanaka, and Hiroshi Nakagawa (Nittobo Acoust. Eng. Co., Ltd. Sumida-ku, Tokyo 130-0021, Japan)

Many parameters exist for evaluating large sound fields such as concert halls. However, it is difficult to apply those parameters for evaluation of a small room such as a recording studio because of their different sound fields. Widely useful common parameters have not been established.

Moreover, early reflections are important in small rooms for determining spatial acoustic impressions. Therefore, various methods that visualize spatial acoustic information obtained by early reflection in rooms are proposed. For this study, sound fields (a music studio and a filmmaking studio) were measured using three kinds of different techniques: instantaneous intensity, mean intensity, and a sphere-baffled microphone array. This report compares the information of sound source directions obtained using these methods. Results show that every method can estimate the position of sound sources and important reflections with high accuracy. In the future, we shall propose a method that visualizes spatial acoustic information more precisely by combining the methods and establishing acoustic parameters that are available for evaluating and designing small rooms.

4:15

**4pAA13. Acoustic evaluation of worship spaces in the city of Curitiba, Brazil.** Cristiane Pulsides, David Q. de Sant'Ana, Samuel Ansay (LAAICA/UFPR, Bloco 4 sala PG-05 81531-990 Curitiba, PR, Brasil, pulsides@gmail.com), Paulo Henrique T. Zannin, and Suzana Damico (LAAICA/UFPR, 81531-990 Curitiba, PR, Brasil)

This article searches acoustic parameters in religious buildings located in the city of Curitiba intending to study its behavior in this kind of facilities. The temples were analyzed according to type of ceremony, architectural style, and construction date. The research was made through the impulsive response integration method for three energetic parameters: (1) reverberation time (RT); (2) clarity (C80); and (3) definition (D50) according recommendations of the ISO/3382:1997 Standard. Performed in between were six and eight impulsive responses in each room using sweep signals and omnidirectional microphones. The results were than compared with referential values already existing [W. Fasold and E. Veres, *Schallschutz + Raumakustik in der Praxis*, 136 (1998)] for acoustic characterizations. It is possible to observe in the measurements the direct connection between reverberation time and the parameters clarity or definition. Moreover, it is possible also to observe the influence of the geometric ratios and architectural elements of the rooms, getting itself for equivalent volumes and rays of removal of the source, different levels of definition.

4:30

**4pAA14. A consideration of the measurement time interval for obtaining a reliable equivalent level of noise from expressway.** Mitsunobu Maruyama (Salesian Polytechnic, Oyamagaoka 4-6-8, Machida, Tokyo 194-0215, Japan) and Toshio Sone (Akita Prepectural Univ., Honjo, Akita 015-0055, Japan)

The level of road traffic noise  $LA_{eq,T}$  greatly depends on the maximum level during the measurement time interval  $\tau$ , and the maximum level often appears at the moment when two consecutive heavy vehicles pass through the point adjacent to the observation point. A mathematical model is proposed for simulating the variation in traffic noise, especially from the point of heavy vehicles with passing. The mean time interval between a pair of two consecutive heavy vehicles with the minimum allowable distance is obtained from time-series data and the mean recurrence time  $h_{ij}$  which can be calculated from the transition matrix  $P[p_{ij}]$ . The comparative study is made among the numbers of heavy vehicles from 25 to 300 [vehicles/hour] in traffic flow and the observation distances of 40 to 200 m from the road. The result shows that the measurement time interval required for the acquisition of reliable data is three to four times as long as  $\tau$  or  $h_{ij}$ .

## Session 4pABa

## Animal Bioacoustics: Marine Mammal Acoustics II

David K. Mellinger, Chair

Oregon State Univ., Hatfield Marine Science Ctr., Newport, OR 97365

## Contributed Papers

1:15

**4pABa1. Great ears: Functional comparisons of land and marine leviathan ears.** D. R. Ketten (Harvard Med. School, Boston, MA; Woods Hole Oceanograph. Inst., Woods Hole, MA), J. Arruda, S. Cramer, M. Yamato (Woods Hole Oceanograph. Inst., Woods Hole, MA), J. O'Malley (Massachusetts Eye and Ear Infirmary, Boston, MA), D. Manoussaki (Vanderbilt Univ., Nashville, TN), E. K. Dimitriadis (NIH/NIDCD, Bethesda, MD), J. Shoshani (Univ. of Asmara, Asmara, Eritrea), and J. Meng (American Museum of Natural History, New York, NY)

Elephants and baleen whales are massive creatures that respond to exceptionally low frequency signals. Although we have many elephant and whale vocalization recordings, little is known about their hearing. Playback experiments suggest hearing in both proboscideans and mysticetes is tuned similarly to low or even infrasonic signals. This raises several interesting issues. First, they emit and perceive signals in two media, air and water, with radically different physical acoustic properties: 4.5-fold differences in sound speed, three-fold magnitude difference in acoustic impedance, and, for common percepts, whales must accommodate 60-fold acoustic pressures. Also, a commonly held tenet is that upper hearing limit is inversely correlated with body mass, implying there should be virtually no whale-elephant hearing overlap given body mass differences. This study analyzed how inner ears in these groups are structured and specialized for low-frequency hearing. Computerized tomography and celloidin histology sections were analyzed in six baleen whale ( $n=15$ ) and two elephant species ( $n=7$ ). The data show mysticetes have a substantially greater hearing range than elephants but that coiling and apical cochlear structures are similar, suggesting common mechanical underpinnings for LF hearing, including cochlear radii consistent with the Whispering Gallery propagation effect. [Work supported by ONR, NIH, WHOI OLI, Seaver Foundation.]

1:30

**4pABa2. Social context of the behavior and vocalizations of the gray whale *Eschrichtius robustus*.** Sarah M. Rohrkasse (School for Field Studies, Ctr. for Coastal Studies, Apartado Postal 15, Puerto San Carlos, BCS, CP 23740 Mexico, sarro101@hotmail.com) and Margaret M. Meserve (Guilford College, Greensboro, NC 27410)

Sound production and surface behavior of the gray whale were investigated at Bahia Magdalena, Mexico to determine if vocalizations have behavioral correlations or are used in specific social contexts. Fifteen-minute sessions of behavioral observations and acoustic recordings of gray whales in various social contexts were collected from February to April 2006 ( $n=30$ ). Analysis of sound production included proportional use of different call types and acoustic variables of each sound type. Preliminary acoustic analysis found no correlation with social contexts or behaviors, but proportional use of different vocalizations is similar to past studies in Baja [Dahlheim *et al*, *The Gray Whale*, pp. 511–541 (1984), F. J. Ollervides, dissertation, Texas A&M University (2001)]. Initial results indicate significant differences in frequencies of high surface behaviors ( $p=0.0477$ ) of groups that include mother-calf pairs. As analysis continues, possible correlations between social context and use of sounds could allow for acoustics to be an indicator of group composition, seasonal movements, and social patterns and to help determine the functions of sounds. [Work supported by SFS and NFWF.]

1:45

**4pABa3. Ambient noise and gray whale *Eschrichtius robustus* behavior.** Francisco Ollervides, Kristin Kuester, Hannah Plekon, Sarah Rohrkasse (School for Field Studies—Ctr. for Coastal Studies, Apartado Postal 15, Puerto San Carlos, BCS, CP 23740 Mexico, follervides@hotmail.com), Kristin Kuester (Univ. of Wisconsin—Madison, Madison, WI 53706), Hannah Plekon (Davidson College, Davidson, NC), and Sarah Rohrkasse (Texas A and M Univ., College Station, TX 77843)

Between 14 February and 13, April 2006, we conducted 31 recording sessions of ambient noise and behavioral sampling of gray whales within Magdalena Bay, Mexico. This breeding lagoon does not have the same Marine Protected Area status compared to the other breeding lagoons of San Ignacio and Guerrero Negro in the Mexican Pacific coast. Poorly monitored guidelines and increasing boat traffic from whale-watching tourism in this area have the potential to affect the surface behavior of these animals and increase average ambient noise levels.

Relative ambient noise levels were recorded and compared to a previous study [Ollervides, 2001] to determine similarities or differences in the 5-year interval between both data sets. Although results are not comparable in decibel levels, probably due to equipment calibration problems, there was a significant difference between the different regions of the bay Kruskal–Wallis ( $p=0.0067$ ). Activity levels ranged from 0.005–0.196 behaviors/whale/minute. Ambient noise levels ranged from 35.70–64.32 dB Re: 1 Pa. No correlation was found between the ambient noise levels in the bay and the activity level of gray whales (correlation value=0.0126; log correlation value=0.172). Further acoustic processing is currently underway.

2:00

**4pABa4. Look who's talking; social communication in migrating humpback whales.** Rebecca A. Dunlop, Michael J. Noad (School of Veterinary Sci., Univ. of Queensland, St. Lucia, Qld 4072, Australia, r.dunlop@uq.edu.au), Douglas H. Cato (Defence Sci. and Tech Org., Pyrmont, NSW 2009, Australia), and Dale Stokes (Scripps Inst. of Oceanogr., La Jolla, CA 92037)

A neglected area of humpback acoustics concerns nonsong vocalizations and surface behaviors known collectively as social sounds. This study describes a portion of the nonsong vocal repertoire and explores the social relevance of individual sound types. A total of 622 different sounds were catalogued and measured from whales migrating along the east coast of Australia. Aural and spectral categorization found 35 different sound types, and discriminate functions supported 33 of these. Vocalizations were analyzed from 60 pods that were tracked visually from land and acoustically using a static hydrophone array. Nonsong vocalizations occurred in all pod compositions: lone whales, adult pairs, mother/calf pairs, mother/calf/escorts, and multiple-adult pods. Thwops and wops were likely to be sex-differentiated calls with wops from females and thwops from males. Sounds similar to song-units were almost all from joining pods and yaps were only heard in splitting pods. Other low-frequency calls (less than 60 Hz) were thought to be within-pod contact calls. Higher-frequency cries (fundamental 450–700 Hz) and other calls (above 700 Hz) and presumed underwater blows were heard more frequently in joining

Pods displaying agonistic behaviors. This work demonstrates that humpbacks produce a great range of contextually different communication signals. [Work supported by ONR and DSTO.]

2:15

**4pABa5. Seasonal ambient noise levels and impacts on communication in the North Atlantic right whale.** Susan E. Parks, Christopher W. Clark, Kathryn A. Cortopassi, and Dimitri Ponirakis (Bioacoustics Res. Program, Cornell Univ., 159 Sapsucker Woods Rd., Ithaca, NY 14850, sep6@cornell.edu)

The North Atlantic right whale is a highly endangered species of baleen whale. Acoustic communication plays an important role in the social behavior of these whales. Right whales are found in coastal waters along the east coast of the United States, an area characterized by high levels of human activity. Most of these activities generate noise that is propagated into the coastal marine environment. The goals of this project are to characterize the noise, both natural and anthropogenic, in right whale habitat areas to determine what levels of noise the whales are regularly exposed to, and whether the acoustic behavior of right whales changes in response to increased noise. Continuous recordings were made from autonomous bottom-mounted recorders in three major habitat areas in 2004 and 2005; Cape Cod Bay (December–May), Great South Channel (May), and the Bay of Fundy, Canada (August) to passively detect right whales by recording their vocalizations. Here, we describe the ambient noise levels in these recordings to describe the daily acoustic environment of right whales, how noise varied over diel, weekly, and seasonal time scales, and whether noise levels correlated with any observed changes in acoustic behavior of the whales.

2:30

**4pABa6. Blue whale calling in Australian waters.** Robert D. McCauley, Chandra P. Salgado Kent (Curtin Univ. of Technol., G.P.O. Box U 1987, Perth 6845, Australia), Christopher L.K. Burton (Western Whale Res. Hillarys 6923, WA Australia), and Curt Jenner (Ctr. for Whale Res. (WA Inc.), Fremantle WA, 6959 Australia)

Calling from the Antarctic true blue whale (*Balaenoptera musculus intermedia*) and the tropical subspecies (*brevicauda*, or pygmy blue) have been recorded across southern Australia with the pygmy blue calls also recorded along the Western Australian (WA) coast. The subspecies have a believed common downsweep and markedly different longer, tonal calls. The frequency of most energy in the tonal calls is offset between the subspecies suggesting sound-space partitioning. The pygmy blue three-part tonal call is typically 120 s long repeated every 200 s, has several variants, and includes a complex two-source component. The nature of the pygmy blue call allows counts of instantaneous calling individuals, giving relative abundance. These estimates in the Perth Canyon, a localized seasonal feeding area, show patterns in usage of space and through time within and between seasons, such as the sudden departure of animals at a season end, which varies by approximately 2 weeks between years. Sea noise records along the WA coast indicate south-traveling animals arrive midway along the coast in October to November, animals fan out across southern Australian over December through May, then move north in the Austral winter. We have begun converting abundance estimates from relative to absolute for pygmy blue calling rates.

2:45

**4pABa7. Acoustical monitoring of finback whale movements on the New Jersey Shelf.** Altan Turgut (Naval Res. Lab., Acoust. Div., Washington, DC 20375) and Christopher Lefler (Univ. of California Santa Barbara, Santa Barbara, CA 93106)

Acoustical monitoring of finback whales is performed by using a data set collected over a 3-week period in December of 2003 on the New Jersey Shelf. One-second-duration 20-Hz signals of finback whales were recorded on three vertical line arrays (VLAs) and a bottomed horizontal line array (HLA). One-second-duration pulses are separated by about 10 s and there is an approximately 2-min-long silent period between 10- to

18-min-long pulse trains. A 30- to 60-min silent period after 5 to 10 pulse trains is also common. Modal analysis of individual pulses indicated that most signals contained two acoustic modes. Arrival-time and group-speed differences of these modes are used for remote acoustic ranging. These modal characteristics are also exploited in a broadband matched-field algorithm for depth discrimination. Bearing estimation of individual whales is obtained by performing horizontal beamforming on the HLA data. Range estimation results are verified by time-of-flight triangulation using single hydrophone data from each VLA location. Acoustic monitoring results indicated that most finback whales traveled near the shelf break front where food might be abundant. Relations between silent periods and acoustic range/depth monitoring results are also investigated. [This work was supported by the ONR.]

3:00–3:15 Break

3:15

**4pABa8. Analysis of melon-headed whale aggregation in Hanalei Bay, July 2004.** David M. Fromm (Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375-5350), Joseph R. Mobley, Jr. (Univ. of Hawaii at M\_noa, Honolulu, HI 96822), Stephen W. Martin (Space and Naval Warfare Systems Ctr. San Diego, San Diego, CA 92152-5001), and Paul E. Nachtigall (Univ. of Hawaii at M\_noa, Kailua, HI 96734)

On 3 July 2004, an aggregation of ca. 150–200 melon-headed whales (*Peponocephala electra*) appeared in the shallow waters of Hanalei Bay, Kauai and congregated there for over 27 h. Preceding the whales' appearance and partially coincident with their time in the Bay, midrange (3.5–5 kHz) tactical sonars were intermittently deployed during the Rim of the Pacific 2004 (RIMPAC) joint military exercises being conducted in waters near Kauai by the U.S., Japan, and Australia Navies. An NOAA report (Southall *et al.*, 2006) attributed the active sonar usage as a plausible, if not likely, contributing factor. A detailed timeline and reconstruction of the RIMPAC activities is presented showing the worst-case estimates of the sonar sound levels in the waters surrounding Kauai. A re-examination of available evidence combined with a new report of a simultaneous and similar aggregation in Sasanhaya Bay, Rota, Commonwealth of the Northern Mariana Islands, brings the plausibility conclusion into question. [This work was sponsored by multiple sources. D. Fromm and S. Martin conducted acoustic analyses with funds provided by the U.S. Pacific Fleet. J. Mobley received funding from the U.S. Geological Survey. P. Nachtigall is sponsored by the Office of Naval Research for marine mammal audiometric studies.]

3:30

**4pABa9. Midfrequency sound propagation in beaked whale environments.** Eryn M. Wezensky, Thomas R. Stottlmyer, Glenn H. Mitchell (Naval Undersea Warfare Ctr., Newport Div., Newport, RI 02841), and Colin D. MacLeod (Univ. of Aberdeen, Aberdeen, U.K.)

Recent mass strandings of beaked whales (*Ziphiidae*, *Cetacea*) coinciding with the use of midfrequency range (1–10 kHz) active sonar have caused speculation about the potentially adverse effects of these sound sources. Particular questions of the research and regulatory communities concern whether beaked whale sensitivity to midfrequency sound exposure is influenced by oceanographic characteristics present at the time of the mass stranding events. This study investigated the interaction between beaked whale habitat characteristics and the nature of a midfrequency signal by analyzing the oceanographic factors affecting underwater acoustic propagation. Three types of model sites were selected from five specific geographical locations where beaked whales have been regularly recorded or where a mass stranding event has been reported. A ray-trace acoustic propagation model was used to generate transmission loss for a 3-kHz signal over a representative 60-km transect at each locality. Model outputs visually demonstrated how the combination of site/event-specific oceanographic characteristics affects the sound propagation of a moving source. A parametric sensitivity comparison and statistical analysis were conducted to identify influential factors between environmental parameters, source depth, and the resulting transmission loss. Major findings of this study as well as future research direction are discussed. [Research supported by NAVSEA.]

**4pABa10. Examination and evaluation of the effects of fast rise-time signals on aquatic animals.** Michael Stocker (Seaflow, Inc., 1062 Ft. Cronkhite, Sausalito, CA 94965)

Increasingly human enterprise is subjecting the ocean environment to acoustic signals to which marine animals are not biologically adapted. This is evidenced by a marked rise in marine mammal strandings, as well as hearing and other physiological damage to fish and other marine organisms as a result of, or coincident to, human-generated noise events. Determining phonotoxic thresholds of marine organisms is complicated by the fact that various marine animals are adapted to sense either pressure gradient or particle motion acoustic energy, or some combination or gradient between the two. This has been addressed to some degree by exposure metrics that consider either net or accumulated acoustical flux densities from various noise sources. This paper examines the role and effects of signal rise time both in terms of physiological impulse response of the exposed organisms, as well as broadband saturation flux densities of fast rise-time signals on animal sense organs. Case studies from the literature will be presented to demonstrate the effects of fast rise time signals on fish. Acoustical signals with high crest factors and fast rise-time components will be compared to signals with dominantly sinusoidal components to illustrate the perceptual effects of these signals on human hearing.

4:00

**4pABa11. Noise separation of underwater acoustic vocalization using auditory filter bank and Poisson rate estimation.** Owen P. Kenny and Craig R. McPherson (Dept. of Elec. and Comput. Eng., James Cook Univ., Douglas 4811, Queensland, Australia)

Formant vocalization tracking has been achieved using a mammalian periphery model and a Poisson rate estimator. This approach used a set of linear bandpass filters to simulate the mechanical displacement of the basilar membrane. The auditory model simulated neural firing by producing a spike at the positive going zero crossing for each filter output. Poisson

intensity of the neural firing rate is controlled by the dominant frequency components of the signal present in the filter. This approach is extended by incorporating neural synchronization information to separate the formant structure from that of noise. The filter structure is designed to overlap the frequency range of adjacent filters. The presence of a formant structure in adjacent filters controls the interspike intervals of neural firing for both filters, which results in the neural firing from both filters being synchronized. If a noise-only component is present in either filter, then the spiking outputs from the adjacent filters are unsynchronized. Experimental results have shown that incorporating neural synchronization information between adjacent filters has enabled separation of signal components from noise. This technique enables easier signal and noise separation than allowed by traditional methods.

4:15

**4pABa12. Using vocalizations of Antarctic seals to determine pupping habitats.** T. L. Rogers, C. J. Hogg, M. B. Ciaglia (Australian Marine Mammal Res. Ctr., Zoological Parks Board of NSW/Faculty of Veterinary Sci., Univ. of Sydney, Mosman Australia), and D. H. Cato (Defence Sci. & Technol. Organisation, Pyrmont, Australia)

The Ross and Leopard seal use the floes of the Antarctic pack ice to whelp and raise their pups, but both species are rarely seen in summer throughout the pack ice. We now realize that this is because they are under the water "calling" during the austral summer as part of their breeding display, and so their presence is underestimated in traditional visual surveys. The period of "calling" overlaps with the time that females give birth, so their vocalizations can be used to determine seal distributions during this time. Acoustic recordings were made using sonobuoys deployed during ship based surveys in the pack ice and analyzed to determine the seal distributions. This was used to predict habitat preference of seals by relating their distributions to remotely sensed indices: ice cover, ice floe type, ice thickness, distance to ice edge, distance to shelf break, distance to land, sea surface temperature, and chlorophyll a.

FRIDAY AFTERNOON, 1 DECEMBER 2006

KOHALA/KONA ROOM, 4:30 TO 5:15 P.M.

## Session 4pABb

### Animal Bioacoustics: Avian Acoustics

Ann E. Bowles, Chair

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

#### Contributed Papers

4:30

**4pABb1. Effective area of acoustic lure surveys for Mexican spotted owls (*Strix occidentalis lucida*).** Samuel L. Denes, Ann E. Bowles (Hubbs-SeaWorld Res. Inst., 2595 Ingraham St., San Diego, CA 92109, sdenes@hswri.org), Kenneth Plotkin, Chris Hobbs (Wyle Labs., Arlington, VA 22202), John Kern (Kern Statistical Services, Sauk Rapids, MN 56379), and Elizabeth Pruitt (GeoMarine, Inc., Hampton, VA 23666)

During acoustic lure surveys for birds, topography and ambient noise are likely to be important determinants of detectability. Examinations of propagation were conducted for acoustic lures (human-made calls) and owl responses recorded during acoustic surveys for Mexican spotted owls in the Gila National Forest (2005). Lure surveys were designed based on

formal agency protocols, which assumed a 0.43-km detection range under typical conditions. A total of 558 points was called over a heavily forested, topographically complex 20x24-km area. Real-time measurements of owl calls and lures were made with a calibrated recording system. Ambient noise was collected using an array of 39 Larson-Davis 820 and 824 sound-level meters. The NMSIM (Wyle Laboratories) single-event propagation simulator was used to model propagation of both owl and human calls. The resulting model of survey effort was compared with a simple two-dimensional statistical model. Probability of detecting owls did not fit the expectations of the agency protocol, suggesting that acoustic propagation should be considered during owl surveys. [Work supported by U.S. Air Force ACC/CEVP; USFWS Permit No. TE024429]

**4pABb2. Automated localization of antbirds and their interactions in a Mexican rainforest.** Alexander N. G. Kirschel, Travis C. Collier, Kung Yao, and Charles E. Taylor (Univ. of California, Los Angeles, 621 Charles E. Young Dr. South, Los Angeles, CA 90095)

Tropical rainforests contain diverse avian communities incorporating species that compete vocally to propagate their signals to intended receivers. In order to effectively communicate with birds of the same species, birds need to organize their song performance temporally and spatially. An automated identification and localization system can provide information on the spatial and temporal arrangement of songs. Acoustic sensor arrays were tested for the ability to localize the source of songs of antbirds recorded in a Mexican rainforest. Pilot studies with a five-node array arranged in a rough circle with a 20-m diameter located the song of Dusky Antbird (*Cercomacra tyrannina*) with an error of 73 cm and Mexican Antthrush (*Formicarius moniliger*) with an error of 65 cm from the location of a source loudspeaker within the array. An additional source 21 m outside was also localized. Results will be presented for experiments and recordings of individuals at the Mexican rainforest site in October 2006. Locations of birds of the same and different species during vocal performance will provide a greater understanding of how individuals interact spatially with each other based on their vocal performance, from which the role of song in ecological interactions can be inferred.

**4pABb3. Nonintrusive acoustic identification of hermit thrush (*Catharus guttatus*) individuals.** Dennis F. Jones (Defence R&D Canada—Atlantic, P.O. Box 1012, Dartmouth, NS, Canada B2Y 3Z7, dennis.jones@drdc-rddc.gc.ca)

From mid-April well into the summer, the secretive hermit thrush (*Catharus guttatus*) can be heard singing throughout the woodlands of Nova Scotia. Its song is distinctive, beginning with a clear introductory note followed by a flurry of flutelike body notes, often cascading and reverberant in character. Despite this fine display of avian virtuosity, few studies have been reported that probe the differences between the calls, songs, and repertoires of individuals. From April 2003 to May 2006, over 3000 songs from several birds were recorded using digital video cameras at study sites in and around the city of Halifax, Nova Scotia. The only birds recorded were those in close proximity to roads and trails. None of the birds were marked, banded, or deliberately disturbed in any way. Although the study birds remained hidden from view most of the time, in the few instances where the birds perched in the open, their behaviors while singing were captured on videotape. All of the birds were readily distinguishable from each other as no two individuals had a single song in common. The most significant finding was that individuals could be reidentified acoustically after 1 week, 3 months, and 1 year had elapsed.

FRIDAY AFTERNOON, 1 DECEMBER 2006

KAHUKU ROOM, 1:00 TO 4:50 P.M.

### Session 4pBB

## Biomedical Ultrasound/Bioresponse to Vibration and Signal Processing in Acoustics: Elastic Imaging

Peter J. Kaczowski, Cochair

*Univ. of Washington, Applied Physics Lab., 1013 NE 40th Street, Seattle, WA 98105-6698*

Tsuyoshi Shiina, Cochair

*Univ. of Tsukuba, Graduate School of Systems and Information Engineering, 1-1-1 Tennodai, Tsukuba 305-8573, Japan*

### Invited Papers

1:00

**4pBB1. Present and future of elasticity imaging technology.** Tsuyoshi Shiina (Grad. School of Systems and Information Eng., Univ. of Tsukuba, 1-1-1 Tennodai Tsukuba, Japan) and Ei Ueno (Univ. of Tsukuba, Tsukuba, Japan)

Elastic properties of tissues are expected to provide us novel diagnostic information since they are based on tissue characteristics and sensitively reflect its pathological state. So far, various techniques for tissue elasticity imaging have been proposed. However, it was not so easy to satisfy real-time operation and freehand manipulation of probe, which was required for practical equipment. To satisfy these conditions, we developed the combined autocorrelation method (CAM) and recently manufactured a commercial ultrasound scanner, for real-time tissue elasticity imaging by implementing the CAM algorithm. By slightly compressing or relaxing the body through freehand operation, the strain images are obtained with real-time and superimposed on B-mode images with a translucent color scale. In addition, we proposed elasticity scores of malignancy by categorizing patterns of elasticity images of breast tumors into five classes from malignant to benign. As a result of diagnosis based on the elasticity score, it was revealed that even nonexperts could attain precise diagnosis of breast cancer based on elasticity score as well as experts since the criterion on elasticity score is much simpler than conventional B-mode images. Finally, some prospects for the next stages of elasticity imaging technology will be surveyed.

1:20

**4pBB2. Real-time tissue elasticity system—Development and clinical application.** Takeshi Matsumura, Tsuyoshi Mitake (Hitachi Medical Corp. 2-1, Toyofuta, Hashiwa-Shi, Chiba-Ken, Japan), Tsuyoshi Tsuyishi, Makoto Yamakawa, Ei Ueno (Tsukuba Univ.), Nobuhiro Fukunari (Shouwa Univ.), and Kumi Tanaka (Nippon Medical Univ.)

The progress of recent semiconductor technology has a remarkable thing. Thanks to progress of this semiconductor technology, the ultrasound scanner in medicine could come to hold enormousness computing power and has come to realize various complicated processing. At the same time, hardness of human tissue which, as you know, is used by palpation is already the information that is important in a diagnosis. But, we think that it does not have enough objectivity. To increase objectivity by visualizing hardness of

tissue, we adopted ECAM (extended combined autocorrelation method), which was developed by Professor Shiina at Tsukuba University in Japan, and succeeded in developing the commercial ultrasound scanner, which could display a strain image in real time. From a clinical point of view, in the breast region, mammography examination is effective in a diagnosis, but a judgment of permeation degree is not superior in ultrasound image. And in a thyroid gland region, we begin to get experience with availability from a diagnosis of papillary cancer and follicular cancer. So, we would like to have the presentation about the development of a strain imaging function and some of our clinical experiences by using the developed system.

1:40

**4pBB3. Elasticity of perfused tissue.** Kirk W. Beach (Dept. of Surgery, Univ. of Washington, Box 356410, Seattle, WA 98195-6410), Barbrina Dunmire, and John C. Kucewicz (Univ. of Washington, Seattle, WA 98105-6698)

Elastic imaging intends to measure Young's modulus (tissue stiffness) or bulk modulus (tissue compressibility) of tissue subjected to an applied strain of several percent. Underlying elastic imaging is the assumption of a linear stress/strain relationship without hysteresis or other time-dependent behavior. Perfused tissue is a composite material comprised of a solid matrix of cells, fibers, interstitial fluid (occupying up to 50% of the tissue volume and varying slowly with time), arterioles (pulsating high-pressure spaces that occupy 0.1% of the tissue volume), capillaries, and venules (low-pressure spaces that occupy up to 3% of the tissue volume varying with respiration). This talk will speculate on the nonlinear, nonstationary stress/strain relationships expected from dependent tissues (legs), pressurized tissues (breast tumors), and other living, perfused tissues. The pressure versus strain curve from each tissue voxel allows the measurement of arteriolar and venular volumes and pressures, and interstitial pressure within the tissues. These volumes and pressures may be key to classifying pathologies.

2:00

**4pBB4. New developments in transient elastography.** Mathias Fink, Mickael Tanter, Ralph Sinkus, and Gabriel Montaldo (LOA, ESPCI, 10 rue Vauquelin, 75005, Paris, France)

An ultra-high-rate ultrasonic scanner has been developed that can give 5000 ultrasonic images per second of the body. With such a high frame rate, the propagation of transient shear waves can be followed, and from the spatio-temporal evolution of the displacement fields, various inversion algorithms allow us to recover the shear modulus map. A discussion on the various inversion algorithms will be presented. In order to obtain unbiased shear elasticity map, different configurations of shear sources induced by radiation pressure of focused transducer arrays are used. Both 2-D and 3-D imaging can be obtained with this technique. *In vitro* and *in vivo* results on breast will be presented that demonstrate the interest of elasticity imaging with transient elastography.

### Contributed Papers

2:20

**4pBB5. Spectral characteristics of breast vibro-acoustography images.** Azra Alizad, Dana H. Whaley, Mathew Urban, Randall R. Kinnick, James F. Greenleaf, and Mostafa Fatemi (Mayo Clinic College of Medicine, Rochester, MN 55905 aza@mayo.edu)

Vibro-acoustography image is a function of the dynamic characteristics of the object at the vibration (difference) frequency ( $df$ ). The dynamic characteristic of tissue is closely related to pathology. Therefore, it is important to evaluate image features versus  $df$ . Here, the influence of  $df$  on breast vibro-acoustography images is studied by scanning human breast at various  $df$  values ranging from 20 to 90 kHz. The subjects were chosen from a group of volunteers with different breast abnormalities. Images were compared subjectively to study image features and the appearances of breast lesions versus  $df$ . It is demonstrated that having a collection of images of the same tissue at different  $df$  values generally provides a better perception of the tissue structure and improves lesion identification. In most cases, higher  $df$  resulted in a higher signal-to-noise ratio and thus a higher image quality. Finally, a frequency-compounded images was obtained by calculating the weighted sum of images at different  $df$  values. It is demonstrated that image compounding normally improves visualization of breast tissue and abnormalities. [Work supported by NIH Grant EB-00535 and Grant BCTR0504550 from the Susan G. Komen Breast Cancer Foundation. Disclosure: Parts of the techniques used here are patented by MF and JFG.]

2:35

**4pBB6. Tissue pulsatility imaging: Ultrasonic measurement of strain due to perfusion.** John C. Kucewicz, Barbrina Dunmire, Lingyun Huang, Marla Paun (Univ. of Washington Appl. Phys. Lab., 1013 NE 40th St., Seattle, WA 98105-6698), and Kirk W. Beach (Univ. of Washington, Seattle, WA 98195-6410)

Over each cardiac cycle perfused tissues expand and relax by a fraction of a percent as blood rapidly accumulates in the arterial vasculature during systole and then slowly drains through the venous vasculature during diastole. Tissue pulsatility imaging (TPI) is a variation on ultrasonic tissue strain imaging that estimates tissue perfusion from this natural, cyclic tissue expansion and relaxation. TPI is derived in principle from plethysmography, a century-old technology for measuring gross tissue volume change from a whole limb or other isolatable body part. With TPI, the plethysmographic signal is measured from hundreds or thousands of sample volumes within an ultrasound image plane to characterize the local perfusion throughout a body part. TPI measures tissue strain over the cardiac cycle and parametrizes the signal in terms of its amplitude and shape. The amplitude of the strain waveform is correlated with perfusion, and the shape of the waveform is correlated with vascular resistance. Results will be presented from the leg showing the change in the TPI signals as the muscles recover from exercise, from breast tumors, and from the brain as blood flow changes in response to visual stimulation. [Work supported in part by NIH 1-R01EB002198-01 and NIH N01-CO-07118.]

4p FRI. PM

**4pBB7. Using human body shear wave noise for passive elastography.**

Karim G. Sabra, Stephane Conti, Philippe Roux, and William A. Kuperman (Scripps Inst. of Ocean., Univ. of California—San Diego, 9500 Gilman Dr., San Diego, CA 92093-0238)

An elastography imaging technique based on passive measurement of shear wave ambient noise generated in the human body (e.g., due to the heart, muscles twitches, and blood flow system) has been developed. This technique merges two recent research developments in medical imaging and physics: (1) recent work on the efficacy of elastographic imaging demonstrating that shear waves are excellent candidates to image tissue elasticity in the human body and (2) theory and experimental verification in ultrasonics, underwater acoustics, and seismology of the concept of extracting coherent Green's function from random noise cross correlations. These results provide a means for coherent passive imaging using only the human body noise field, without the use of external active sources. Coherent arrivals of the cross correlations of recordings of human body noise in the frequency band 2–50 Hz using skin-mounted accelerometers allows us to estimate the local shear velocity of the tissues. The coherent arrivals emerge from a correlation process that accumulates contributions over time from noise sources whose propagation paths pass through both sensors. The application of this passive elastography technique for constructing biomechanical models of *in vivo* muscles' properties will be discussed.

**3:05–3:20 Break****3:20**

**4pBB8. Dynamic radiation force of acoustic waves on solid elastic spheres.** Glauber T. Silva (Instituto de Computação, Universidade Federal de Alagoas, Maceió, AL, 57072-970, Brazil)

The present study concerns the dynamic radiation force on solid elastic spheres exerted by a plane wave with two frequencies (bichromatic wave) considering the nonlinearity of the fluid. Our approach is based on solving the wave scattering for the sphere in the quasilinear approximation within the preshock wave range. The dynamic radiation force is then obtained by integrating the component of the momentum flux tensor at the difference of the primary frequencies over the boundary of the sphere. Effects of the fluid nonlinearity play a major role in dynamic radiation force, leading it to a regime of parametric amplification. The developed theory is used to calculate the dynamic radiation force on three different solid spheres (aluminum, silver, and tungsten). The obtained spectrum of dynamic radiation force presents resonances with larger amplitude and better shape than those exhibited in static radiation force. Applications of the results to some elasticity imaging techniques based on dynamic radiation force will be presented.

**3:35**

**4pBB9. Ultrasonic measurement of displacement distribution inside an object caused by dual acoustic radiation force for evaluation of muscular relax property due to acupuncture therapy.** Yoshitaka Odagiri, Hideyuki Hasegawa, and Hiroshi Kanai (Grad. School of Eng., Tohoku Univ., Sendai 980-8579, Japan, odagiri@us.ecei.tohoku.ac.jp)

Many studies have been carried out on the measurement of mechanical properties of tissues by applying an ultrasound-induced acoustic radiation force. To assess mechanical properties, strain of an object must be generated. However, one radiation force is not sufficient because it also causes translational motion when the object is much harder than surrounding medium. In this study, two cyclic radiation forces are applied to a muscle phantom from two opposite horizontal directions so that the object is

cyclically compressed in the horizontal direction. As a result, the object is vertically expanded due to the incompressibility. The resultant vertical displacement is measured using ultrasound. Two concave ultrasonic transducers for actuation were both driven by sums of two continuous sinusoidal signals at two slightly different frequencies of 1 MHz and (1M + 5) Hz. Displacement, which fluctuates at 5 Hz, was measured by the ultrasonic *phased tracking method* proposed by our group. Results indicated that the surface of the phantom was cyclically actuated with an amplitude of a tenth of a few micrometers, which well coincided with that measured with laser vibrometer. In addition, upward and downward displacements at the surface and deeper region were found during the increase phase of radiation forces. Such displacements correspond to the horizontal compression.

**3:50**

**4pBB10. A phantom study on ultrasonic measurement of arterial wall strain combined with tracking of translational motion.** Hideyuki Hasegawa and Hiroshi Kanai (Grad. School of Eng., Tohoku Univ., Aramaki-aza-Aoba 6-6-05, Sendai 980-8579, Japan, hasegawa@us.ecei.tohoku.ac.jp)

Correlation-based techniques are often applied to ultrasonic rf echoes to obtain the arterial wall deformation (strain). In such methods, the displacement estimates are biased due to changes in center frequency of echoes. One of the reasons for the change in center frequency is the interference of echoes from scatterers within the wall. In the phased tracking method previously proposed for strain estimation by our group, the estimated displacement contains both the components due to the translational motion and strain. The translational motion is larger than strain by a factor of 10 and, thus, the error in the estimated displacement due to the change in center frequency mainly depends on translational motion and is often larger than the minute displacement due to strain. To reduce this error, in this study, a method is proposed in which the translational motion is compensated using the displacement of the luminal boundary estimated by the phased tracking method before correlating echoes between the frame before deformation and that at the maximum deformation to estimate the strain distribution within the wall. In basic experiments using phantoms made of silicone rubber, the estimation error was much reduced to 15.6% in comparison with 36.4% obtained by the previous method.

**4:05**

**4pBB11. Wave biomechanics of skeletal muscle.** Oleg Rudenko (Dept. of. Blekinge Inst. of Technol., 371 79 Karlskrona, Sweden) and Armen Sarvazyan (Artann Labs., Inc., West Trenton, NJ 08618)

Physiological functions of skeletal muscle, such as voluntary contraction and force development, are accompanied by dramatic changes of its mechanical and acoustical properties. Experimental data show that during contraction, the muscle's Young's modulus, shear viscosity, and anisotropy parameter are changed by over an order of magnitude. None of the existing models of muscle contraction and muscle biomechanics can adequately explain the phenomena observed. A new mathematical model [O. Rudenko and A. Sarvazyan, *Acoust. Phys.* (6), (2006)], has been developed relating the shear wave propagation parameters to molecular structure of the muscle and to the kinetics of the mechanochemical cross-bridges between the actin and myosin filaments. New analytical solutions describing waves in muscle including nonlinear phenomena are found. A molecular mechanism for the dependence of acoustical characteristics of muscle on its fiber orientation and the contractile state is proposed. It is shown that although the anisotropy connected with the preferential direction along the muscle fibers is characterized by five elastic moduli, only two of these moduli have independent values in the muscle. The potential implications of the proposed model in terms of the acoustical assessment of muscle function are explored.

**4pBB12. Phase aberration correction for a linear array transducer using ultrasound radiation force and vibrometry optimization: Simulation study.** Matthew W. Urban and James F. Greenleaf (Dept. of Physiol. and Biomed. Eng., Mayo Clinic College of Medicine, 200 First St. SW, Rochester, MN 55905, urban.matthew@mayo.edu)

Diagnostic ultrasound images suffer from degradation due to tissues with sound speed inhomogeneities causing phase shifts of propagating waves. These phase shifts defocus the ultrasound beam, reducing spatial resolution and image contrast in the resulting image. We describe a phase aberration correction method that uses dynamic ultrasound radiation force to harmonically excite a medium using amplitude-modulated continuous wave ultrasound created by summing two ultrasound frequencies at  $f_0 = 3.0$  MHz and  $f_0 + \Delta f = 3.0005$  MHz. The phase of each element of a linear array transducer is sequentially adjusted to maximize the radiation force and obtain optimal focus of the ultrasound beam. The optimization is performed by monitoring the harmonic amplitude of the scatterer velocity in the desired focal region using Doppler techniques. Simulation results show the ability to regain a 3.0-MHz focused field after applying a phase screen with an rms time delay of 95.4 ns. The radiation force magnitude increased by 22 dB and the resolution of the field was regained. Simulation results show that the focus of the beam can be qualitatively and quantitatively improved with this method. [This study was supported in part by Grants EB002640 and EB002167 from the NIH.]

**4pBB13. Application of the optoacoustic technique to visualization of lesions induced by high-intensity focused ultrasound.** Tatiana Khokhlova, Ivan Pelivanov, Vladimir Solomatn, Alexander Karabutov (Intl. Laser Ctr., Moscow State Univ., 119992, Moscow, Russia t\_khokhlova@ilc.edu.ru), and Oleg Sapozhnikov (Moscow State Univ., 119992, Moscow, Russia)

Today several techniques are being applied to monitoring of high-intensity focused ultrasound (HIFU) therapy, including MRI, conventional ultrasound, and elastography. In this work a new method for noninvasive monitoring of HIFU therapy is proposed: the optoacoustic method. The optoacoustic technique is based on the excitation of wideband ultrasonic pulses through the absorption of pulsed laser radiation in tissue and subsequent expansion of the heated volume. The excited optoacoustic (OA) pulse contains information on the distribution of optical properties within the tissue—light scattering and absorption coefficients. Therefore, if thermal lesions have different optical properties than the untreated tissue, they will be detectable on the OA waveform. The considerable change in light scattering and absorption coefficients after tissue coagulation was measured using techniques previously developed by our group. Heating induced by HIFU also influences the OA signal waveform due to the rise of thermal expansion coefficient of tissue with temperature. This dependence was measured in order to evaluate the feasibility of the OA technique in temperature monitoring. An OA image of HIFU lesion induced by a 1.1 MHz focused transducer in a liver sample was reconstructed using a 64-element wideband array transducer for OA signal detection.

FRIDAY AFTERNOON, 1 DECEMBER 2006

OAHU ROOM, 1:00 TO 3:00 P.M.

### Session 4pEAa

## Engineering Acoustics: New Electroacoustic Transducers Utilizing Advanced Technologies and Materials

Juro Ohga, Cochair

*Shibaura Inst. of Technology, 3-9-14 Shibaura, Minato-ku, Tokyo 108-8548, Japan*

James E. West, Cochair

*Johns Hopkins Univ., Dept. of Electrical and Computer Engineering, Barton 105, 3400 N. Charles St., Baltimore, MD 21218-2686*

### Invited Papers

1:00

**4pEAa1. Solid-state photo-microphones or pressure sensors by total reflection.** Yasushi Suzuki (Gunma Natl. College of Tech., 580, Toriba-cho, Maebashi-shi, Gunma, 371-8530 Japan., suzuki@elc.gunma-ct.ac.jp) and Ken'iti Kido (Tohoku Univ., Yokohama-shi, Kanagawa, 226-0017 Japan)

Solid-state photo-microphones or pressure sensors are proposed. These sensors use a new principle, involving the optical total reflection at the boundary surface between glass and air. The critical angle for total reflection changes by the refractive index of air, which depends on the air density. Sound pressure changes the air density. Therefore, the sound pressure is measurable by detecting the intensity of the reflected light from the total reflection area. The sensitivity of the sensor is investigated theoretically. It is expected that the sensor has sufficient sensitivity for practical use, employing laser light and a curved boundary surface with a large radius of curvature. Some experiments are carried out to verify the theoretical investigations. A He-Ne laser or a laser diode is employed as a light source in the experiments. Experimental results show that the sensor has equivalent sensitivity to that which was theoretically estimated, but that sensitivity is very low. The sensor is useful as a pressure sensor, but it is difficult to realize a microphone for general use at the present. The microphones have no diaphragm and the upper limit in the frequency range is extremely high in principle.

**4pEAa2. Micromachined microphones with diffraction-based optical interferometric readout.** F. Levent Degertekin (G.W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332, levent@gatech.edu), Neal A. Hall (Sandia Natl. Labs, Albuquerque, NM 87185-5800), and Baris Bicen (Georgia Inst. of Technol., Atlanta, GA 30332)

A diffraction-based optical method for integrated interferometric detection of micromachined microphone diaphragm displacement is described. With multichip optoelectronics integration, this approach yields highly sensitive optical microphones in mm-cube volumes. Since the microphone sensitivity does not depend on capacitance, this method changes the paradigm for the backplate and gap structure design. As a result, one can use millimeter size diaphragms to achieve wide frequency response and low thermal mechanical noise levels characteristic of precision measurement microphones. Furthermore, the electrical port of the device, which is freed by optical detection, is used for electrostatic actuation of the microphone diaphragm to tune microphone sensitivity and to generate self-characterization signals. Prototype optical microphone structures have been fabricated using Sandia National Laboratories' silicon based SwIFT-Lite™ process. Measurements on these diaphragms show an A-weighted diaphragm displacement noise of 2.4 pm and flat electrostatic response up to 20 kHz. These results indicate the feasibility of realizing measurement microphones with 1.5-mm-diam diaphragms, 15-dBA internal noise, and 40-kHz bandwidth. Application of the detection method in a bio-inspired directional microphone for hearing aids is also discussed. [Work partially supported by NIH Grant 5R01DC005762-03, Sensing and Processing for Hearing Aids.]

**4pEAa3. Hardware and software technologies for improvement of hearing characteristics of headphone reproduction.** Kiyofumi Inanaga and Yuji Yamada (Audio Codec Development Dept., Technol. Development Group, SONY Corp., Shinagawa Tec., 12-15-3, Tokyo, 108-6201 Japan)

This report specifically describes commercialization technology of a headphone system with out-of-head localization applying dynamic head-related transfer functions (HRTFs) that can localize sound easily over a full 360 deg. A source image by output of conventional headphones is localized inside the listener's head. However, the image can be localized outside the listener's head by wearing headphones over a full 360 deg through accurate simulation of the listener's HRTFs. Developments of headphone systems using signal processing technology for data correction have given rise to the static binaural reproduction system (SBRS). The first part of this speech describes its psychoacoustic characteristics and challenges. A rotating dummy-head that is synchronized with the listener's head movement was produced experimentally to create the dynamic binaural reproduction system (DBRS). Using the DBRS, HRTFs synchronize with the listener's head movement. Psychoacoustic characteristics and advantages of the system are also discussed in this report. Further developments were made to realize the commercialization of the DBRS in areas including piezoelectric gyroscope head-tracking technology, headphone technologies that can reproduce real sound characteristics, and simplification of HRTF signal processing employing a simulator with electronic circuits. Finally, future visions for these technologies will be touched upon.

**4pEAa4. Piezoelectret microphones: A new and promising group of transducers.** Gerhard M. Sessler and Joachim Hillenbrand (Darmstadt Univ. of Technol., Merckstrasse 25, 64283 Darmstadt, Germany, g.sessler@nt.tu-darmstadt.de)

Piezoelectret microphones, first described a few years ago, are transducers based on the strong longitudinal piezoelectric effect of charged cellular polymers. Such microphones have recently been improved in two respects: Firstly, an expansion process was used to increase the piezoelectric  $d_{33}$  coefficients of cellular polypropylene (PP) films in the audio frequency range up to 600 pC/N and, secondly, stacking of several films was applied to increase the microphone sensitivity. Transducers with six films now show open-circuit sensitivities of up to 15 mV/Pa, comparable to that of electret microphones. Other characteristics of piezoelectret microphones are their low equivalent noise level of about 26 dB(A) and the very small total harmonic distortion of less than 0.1% at 140 dB SPL. The piezoelectric activity of the PP films and the microphone sensitivities are stable at room temperature but start to decay above 50 °C. Recently, directional piezoelectret microphones with various directional characteristics have been designed. Major advantages of piezoelectret microphones are their simple design, their low harmonic distortion, and their wide frequency range extending into the ultrasonic region.

**4pEAa5. Expansion of frequency range for piezoelectric loudspeakers by new transducer construction.** Juro Ohga (Shibaura Inst. of Technol., 3-7-5, Toyosu, Koto-ku, Tokyo 135-8548, Japan)

Although simple construction of piezoelectric loudspeakers engenders various merits, expansion of its working frequency range to the very low region is difficult because the mechanically stiff characteristics of conventional piezoelectric ceramic diaphragms prevent their large amplitude operation. This paper proposes two sorts of new piezoelectric loudspeaker construction that are suitable for low-frequency signal radiation. One idea is the use of a tuck-shape diaphragm by a PVDF polymer film bimorph. It has large surface area with a very low resonant frequency. Resonant frequencies and sensitivity frequency characteristics are examined, and control methods of local diaphragm bending are discussed. The other idea is the use of continuous revolution of a piezoelectric ultrasonic motor. It produces a completely controlled large output force because its output mechanical impedance is much greater than that of any conventional transducer or motor. An ultrasonic motor, whose stator is connected to a direct-radiator loudspeaker cone by a rod and whose rotor is burdened by a heavy metal ring, rotates with a constant velocity. Modulation of the velocity by using an audio signal imparts a driving force to the diaphragm because the heavy ring tends to keep a constant velocity. Experimental models suggest that this construction is useful.

**4pEAa6. Modal array signal processing using circular microphone arrays applied to acoustic source detection and localization problems.** Heinz Teutsch (Avaya Labs, 233 Mt. Airy Rd., Basking Ridge, NJ 07920, teutsch@avaya.com) and Walter Kellermann (Univ. of Erlangen-Nuremberg, Erlangen, Germany)

Many applications of acoustic signal processing rely on estimates of several parameters present in the observed acoustic scene such as the number and location of acoustic sources. These parameters have been traditionally estimated by means of classical array signal processing (CASP) algorithms using microphone arrays. Algorithms for parameter estimation solely based on the paradigm of CASP often suffer from the narrowband assumption underlying the signal model. This restriction limits their usability when wideband signals, such as speech, are present in the wave field under observation. We investigate the parameter estimation problem by applying the notion of wave field decomposition using baffled circular microphone arrays. The obtained wave field representation is used as the basis for "modal array signal processing algorithms." It is shown that by applying the notion of modal array signal processing, novel algorithms can be derived that have the potential to unambiguously detect and localize multiple simultaneously active wideband sources in the array's full field-of-view. Performance evaluations by means of simulations, measurements, and real-time case studies are presented.

FRIDAY AFTERNOON, 1 DECEMBER 2006

OAHU ROOM, 3:15 TO 6:00 P.M.

### Session 4pEAb

## Engineering Acoustics: Special Topics in Engineering Acoustics

Timothy W. Leishman, Cochair

*Brigham Young Univ., Dept. of Physics and Astronomy, N247 ESC, Provo, UT 84602*

Kiyofumi Inanaga, Cochair

*Sony Corp., Shinagawa Tec. 12-15-3, Tokyo 108-6201, Japan*

### Contributed Papers

3:15

**4pEAb1. Enhanced voided piezoelectric polymer for underwater acoustic sensors.** Juan Arvelo (Appl. Phys. Lab., Johns Hopkins Univ., 11100 Johns Hopkins Rd., Laurel, MD 20723-6099), Ilene Busch-Vishniac, and James West (Johns Hopkins Univ., Baltimore, MD 21218)

A charged voided polymer has been shown to exhibit large piezoelectricity. This material consists of injected air bubbles into polypropylene. This sheet of voided material is then biaxially stretched to elongate the voids. After stretching this material, a strong electric field is applied to cause dielectric breakdown of the gas in the voids, creating electric charges that are trapped in the polymer frame. Since the sides of the voids have opposite charges, they form macroscopic dipoles. When an external force is applied to this material, the voids become narrower, causing stronger dipole strength. A simple model of this voided material was implemented to derive formulas to estimate its piezoelectric constant, electro-mechanical coupling factor, resonance frequency, and sensor sensitivity based on electrical and mechanical properties of the polymer and gas in the voids. These formulas and a survey of available polymers and gases yielded promising combinations that result in more sensitive voided materials that satisfy selected criteria. These criteria include high sensitivity and maximum service temperature, low dissipation factor, and high dynamic compressibility, but low hydrostatic compressibility. This talk will describe the model, derive the formulas, uncover measured properties of candidate polymers and gases, and show calculated sensitivity of selected polymer/gas combinations.

3:30

**4pEAb2. Basic study on one-dimensional transducer array using hydrothermally synthesized lead zirconium titanate poly-crystalline film.** Akito Endo, Tomohito Hasegawa, Norimichi Kawashima, Shinichi Takeuchi (1614, Kurogane-cho, Aoba-ku, Yokohama, Kanagawa, 225-8502, Japan), Mutsuo Ishikawa, and Minoru Kurosawa (Midori-ku, Yokohama, Kanagawa 226-8502, Japan)

Recently, high-frequency miniature medical ultrasound probes with high resolution were actively developed. However, it is difficult to fabricate such tiny ultrasound probes using piezoelectric ceramic vibrator with thickness less than 100  $\mu\text{m}$ . We deposited a PZT poly-crystalline film on a titanium substrate using the hydrothermal method and developed transducers using the PZT poly-crystalline film for ultrasound probes. In this study, we applied it to a miniature medical one-dimensional (1-D)-array-type ultrasound probe with resonance frequency of 10 MHz. After sputtering of pure titanium on the surface of a hydroxyapatite substrate, the titanium film was etched using the photolithography method to form a 1-D titanium film electrode array with 75  $\mu\text{m}$  element pitch, 40  $\mu\text{m}$  element width, and 4 mm element length to scan an ultrasound beam electronically by sector scan mode using phased-array technique. Thereby we fabricated a miniature 1-D-array-type ultrasound probe. A transmitted ultrasound pulse from 10 MHz commercial ultrasound probe was received by this fabricated 1-D-array type ultrasound probe with hydrothermally synthesized PZT poly-crystalline film vibrators.

**4pEAb3. Analysis of a barrel-stave flextensional transducer using MAVART (model to analyze the vibrations and acoustic radiation of transducers) and ATILA (analysis of transducers by integration of Laplace equations) finite-element codes.** Richard A. G. Fleming, Mark Kwiecinski, and Dennis F. Jones (Defence R&D Canada—Atlantic, P.O. Box 1012, Dartmouth, NS, Canada B2Y 3Z7, dennis.jones@drdc-rddc.gc.ca)

A small barrel-stave flextensional transducer, designed and tested at Defence Research and Development Canada—Atlantic, is a candidate sound source for underwater coastal surveillance and acoustic communications applications. This high-power transducer has an outside diameter, length, and mass of 5.7 cm, 12.7 cm, and 1.1 kg, respectively. The measured fundamental flexural resonance frequency was 1.8 kHz with a transmitting voltage response of 118 dB/1 $\mu$ Pa-m/V and an omnidirectional radiation pattern. Two finite-element models were developed for this transducer using the finite-element codes MAVART (Model to Analyze the Vibrations and Acoustic Radiation of Transducers) and ATILA (Analysis of Transducers by Integration of Laplace equations). Comparisons are made between the calibration measurements and the model predictions. [Work supported in part by Sensor Technology Limited.]

4:00

**4pEAb4. Thermal behavior of high-power active devices with the ATILA (analysis of transducers by integration of Laplace equations) finite-element code.** Jean-Claude Debus (Institut Supérieur de l'Electronique et du Numerique, 41 Bv Vauban, 59046 Lille, Cedex France), John Blottman III, and Stephen Butler (Naval Undersea Warfare Ctr. Div. Newport, RI 02841)

Many active devices using piezoelectric ceramics are driven with very high power densities and long pulse lengths. Due to mechanical and dielectric losses in the materials, this produces heat, causing a temperature rise in the devices, which may lead to their mechanical failure. The thermal issues have been shown to be the limiting device design criteria over electric field and mechanical stress limits, yet the effect of the temperature on performance is generally not considered in the numerical models used during the design stage. A coupled electro-mechanical thermal analysis is implemented in the ATILA code. For a steady-state or transient solution, a thermal behavior is weakly coupled to the electromechanical response. The method may take advantage of the order-of-magnitude-greater time constant for thermal effects compared to mechanical behavior. A two-step analysis is performed whereby the electromechanical behavior is first computed, and the resulting dissipated power is then applied as a heat generator to determine the resulting temperature of the device. A high-drive, 31-mode, free flooded ring transducer and a sonar projector serve as validation of the numerical model. The approach addresses both the transient thermal response and the steady temperature profile that results from the high-power, high-duty-cycle drive.

4:15

**4pEAb5. Development of multichannel optical sensor and visualization of vibration distribution.** Jun Hasegawa and Kenji Kobayashi (Faculty of Eng., Takushoku Univ., 815-1 Tatemachi, Hachioji-shi, Tokyo 193-0985 Japan, jhase@es.takushoku-u.ac.jp)

A multi-channel optical sensor system was developed to measure vibrations simultaneously with high spatial resolution. As sensor elements, optical displacement sensor units were developed not to disturb the natural vibration. Each sensor unit, which consists of the optical fiber bundle and focusing lens, can detect the displacement of the object as the variation of the reflected light power. The sensor unit has a displacement resolution of 10 nm, a dynamic range of more than 90 dB, and a frequency band width of up to 80 kHz. Up to 64 sensor units can be arrayed as one sensor head, which realizes the simultaneous measurement of vibration distribution with the high spatial resolution of 4 mm. A calibrating function under the measurement circumstances was developed. Under calibration mode, the sensor array head is moved by a linear actuator, while the vibration of the object is stopped. Thus the calibrated data of each sensor unit can be

obtained for the displacement magnitude. Measured vibration distributions can be monitored as the three-dimensional animations. With the system developed, several actuators for vibratory micro-injection were measured, and the system could reveal their detailed vibration distributions and could detect the existence of a failure portion of some actuator.

4:30

**4pEAb6. Prediction of howling for a sound system in an acoustical environment with both reverberant and direct sounds.** Hideki Akiyama and Juro Ohga (Shibaura Inst. of Technol., 3-7-5 Toyosu, Koto-ku, Tokyo 135-8548, Japan, m106003@shibaura-it.ac.jp)

Prediction of howling is a key technology for a howling suppression design for a sound system with a loudspeaker and microphone. A howling occurrence prediction method for a sound system in a reverberant room has already been presented [J. Ohga, J. Sakaguchi, "Prediction of howling of a sound system in a reverberant room," W. C. Sabine Centennial Symposium (ASA, New York, 1994), 2aAAd4]. It is apparently useful for ordinary public address systems whose distances of loudspeakers from microphones are large. However, this result was not perfect because the direct sound component is not negligible in hands-free telephones or tele-conference systems whose loudspeakers and microphones are set close to each other. This report gives a quantitative howling occurrence prediction method for a sound system in an acoustical environment with both reverberant and direct sounds. The following design parameters are obtained: (1) the increase of howling occurrence level from the power average value, (2) the level occurrence probability, and (3) the critical level chart given by an equation as a function of direct and reverberant sounds ratio. Prediction results for particular examples are compared with calculations of sound-field transfer functions. Results confirmed that it is practical.

4:45

**4pEAb7. Effect of background noise on dialogue in telephony.** Koichi Amamoto and Juro Ohga (Sibaura Inst. of Technol., 3-7-5 Toyosu, Koto-ku, Tokyo, 135-8548, Japan, m106006@shibaura-it.ac.jp)

Recent developments of mobile telephones include new sorts of impairments against speech. Conventional evaluation method for impairments by a talker and a few listeners cannot apply to these new ones, because they are brought by long signal delay. The effect of it cannot discriminate by "one-sided" test. This research relates to a speech quality evaluation by conversation between two persons. Variation of conversation stream is observed by addition of pink noise of various levels to a dialogue by microphones and earphones. Length of sentences and frequency of repeats are quantified and their meanings are discussed

5:00

**4pEAb8. Best practices for auditory alarm design in space applications.** Durand Begault and Martine Godfroy (Human Systems Integration Div., NASA Ames Res. Ctr., Moffett Field, CA 94035)

This presentation reviews current knowledge in the design of auditory caution and warning signals, and sets criteria for development of "best practices" for designing new signals for NASA's Crew Exploration Vehicle (CEV) and other future spacecraft, as well as for extra-vehicular operations. A design approach is presented that is based upon cross-disciplinary examination of psychoacoustic research, human factors experience, aerospace practices, and acoustical engineering requirements. Existing alarms currently in use with the NASA Space Shuttle flight deck are analyzed and then alternative designs are proposed that are compliant with ISO 7731 ("Danger signals for work places Auditory Danger Signals") and that correspond to suggested methods in the literature to insure discrimination and audibility. Parallel analyses are shown for a sampling of medical equipment used in surgical, periop, and ICU contexts. Future development of auditory sonification techniques into the design of alarms will allow auditory signals to be extremely subtle, yet extremely useful in subtly indicating trends or root causes of failures. [Work funded by NASA's Space Human Factors Engineering Project.]

5:15

**4pEAb9. Acoustic signal analysis for forensic applications.** Durand Begault and Christopher Peltier (Audio Forensic Ctr., Charles M. Salter Assoc., 130 Sutter St., Ste. 500, San Francisco, CA 94104, durand.begault@cmsalter.com)

Acoustical analysis of audio signals is important in many legal contexts for determining the authenticity, originality, and continuity of recorded media; determining the circumstances of events in question that may have been recorded; for determining the audibility of signals; and for identification or elimination of talkers as a match to an unknown exemplar. Recorded media are analyzed in forensic applications using both familiar techniques (waveform and spectral analyses) and more novel methods (e.g., ferro fluid development of media; specialized tape heads with non-standard reproduction characteristics; crystal microscopy; detection and matching to power grid frequencies). Audibility analyses frequently require careful reconstructive field measurements and criteria in excess of normally accepted standards. Voice identification-elimination protocols must account for examiner bias and exemplar quality and can be described using a receiver operator curve (ROC) model. This presentation gives an overview of these techniques and their comparative advantages for specific forensic applications.

5:30

**4pEAb10. Without low-pass filter for the 1-bit digital amplifier.** Kiyoshi Masuda (Coroprate Res. and Development Group, SHARP Corp., 2613-1 Ichinomoto-cho, Tenri-shi, Nara, Japan) and Yoshio Yamasaki (Waseda Univerity, Okubo, Shinzyuku-ku, Tokyo, Japan)

SHARP collaborated with Waseda University from 1990 for 1-bit digital technology. SHARP had started to receive an order for the 1-bit digital amplifier "SM-SX100" on 20 August 1999. Until today, we have introduced the 1-bit digital amplifier for audio, flat panel TV (LCD TV), and PC. These 1-bit amplifiers provided low-pass filter for the final stage,

which is provided after 1-bit digital switching. We have to achieve more good sound and reduce deterioration of this low-pass filter. We have introduced new 1-bit digital amplifier without this low-pass filter beginning this April. This means we controlled the 1-bit digital signal to directly operate the speaker. We have proved a better effect for sound to compare the new 1-bit digital amplifier with the PWM switching amplifier, the A-class amplifier and the 1-bit digital amplifier with low-pass filter. If we do not measure any improvement for this new 1-bit digital amplifier, it has large radiation noise. We had achieved a reduction to the limit level of FCC, Denanhou, etc.

5:45

**4pEAb11. Force-frequency effect of thickness mode langasite resonators.** Haifeng Zhang (W317.4 Nebraska Hall, Univ. of Nebraska, Lincoln, NE 68588-0526, hfzhang@bigred.unl.edu), Joseph A. Turner, Jiashi Yang (Univ. of Nebraska, Lincoln, NE 68588-0526), and John A. Kosinski (U.S. Army CECOM, Fort Monmouth, NJ 07703-5211)

Langasite resonators are of recent interest for a variety of applications because of their good temperature behavior, good piezoelectric coupling, low acoustic loss, and high Q factor. The force-frequency effect describes the shift in resonant frequency a resonator experiences due to the application of a mechanical load. A clear understanding of this effect is essential for many design applications such as pressure sensors. In this presentation, the frequency shift is analyzed theoretically and numerically for thin, circular langasite plates subjected to a diametrical force. The results are compared with experimental measurements of the same system for a variety of langasite resonators with various material orientations. In addition, the sensitivity of force-frequency effect is analyzed with respect to the nonlinear material constants. A comparison between the force-frequency effect of langasite and quartz resonators is also made. Finally, the application of such measurements for determining third-order elastic constants is discussed. [Work supported by ARO.]

FRIDAY AFTERNOON, 1 DECEMBER 2006 IAO NEEDLE/AKAKA FALLS ROOM, 1:20 TO 4:25 P.M.

### Session 4pMU

## Musical Acoustics and Psychological and Physiological Acoustics: Acoustic Correlates of Timbre in Musical Instruments

James W. Beauchamp, Cochair

*Univ. of Illinois Urbana-Champaign, School of Music, Dept. of Electrical and Computer Engineering, 1002 Eliot Dr., Urbana, IL 61801*

Mashashi Yamada, Cochair

*Kanazawa Inst. of Technology, Dept. of Media Informatics, 3-1 Yatsukaho, Hakusan, Ishikawa 924-0838, Japan*

### Invited Papers

1:20

**4pMU1. A meta-analysis of acoustic correlates of timbre dimensions.** Stephen McAdams, Bruno Giordano (CIRMMT, Schulich School of Music, McGill Univ., 555 Sherbrooke St. West, Montreal, QC, Canada H3A 1E3), Patrick Susini, Geoffroy Peeters (STMS-IRCAM-CNRS, F-75004 Paris, France), and Vincent Rioux (Maison des Arts Urbains, F-75020 Paris, France)

A meta-analysis of ten published timbre spaces was conducted using multidimensional scaling analyses (CLASCAL) of dissimilarity ratings on recorded, resynthesized, or synthesized musical instrument tones. A set of signal descriptors derived from the tones was drawn from a large set developed at IRCAM, including parameters derived from the long-term amplitude spectrum (slope, centroid, spread, deviation, skewness, kurtosis), from the waveform and amplitude envelope (attack time, fluctuation, roughness), and from variations in the short-term amplitude spectrum (flux). Relations among all descriptors across the 128 sounds were used to determine families of related descriptors and to reduce the number of descriptors tested as predictors. Subsequently multiple correlations between descriptors and the positions of timbres along perceptual dimensions determined by the CLASCAL analyses were

computed. The aim was (1) to select the subset of acoustic descriptors (or their linear combinations) that provided the most generalizable prediction of timbral relations and (2) to provide a signal-based model of timbral description for musical instrument tones. Four primary classes of descriptors emerge: spectral centroid, spectral spread, spectral deviation, and temporal envelope (effective duration/attack time). [Work supported by CRC, CFI, NSERC, CUIDADO European Project.]

1:40

**4pMU2. Perceptual acoustics of consonance and dissonance in multitimbral triads.** Roger Kendall (Music Cognition and Acoust. Lab., Program in Systematic Musicology, UCLA, 405 Hilgard Ave., Los Angeles, CA 90095, kendall@ucla.edu) and Pantelis Vassilakis (DePaul Univ., Chicago, IL 60614)

Most studies of consonance and dissonance assume a singular spectrum for the constituent intervals of a dyad. Recently, the principal author conducted experiments evaluating triads consisting of digitally mixed combinations drawn from the MUMS single-note natural-instrument recordings. Results indicated that the main effect of ratings for consonance and dissonance correlated well with studies using artificial signals. However, interaction effects suggested perceptual differences related to the timbral differences across combinations. The present experiment evaluates perceptual and acoustical variables of the ten possible triadic combinations created with C4 as the lower and the ten with C5 as the upper notes. UCLA wind ensemble performers on oboe, flute, and clarinet, combinations designed to span timbral space, were digitally recorded. Analyses include perceptual ratings of consonance and dissonance, similarity, as well as acoustical analysis of roughness using a recently developed model. Since natural performances of any type vary in fundamental frequency, additional experiments will employ emulated oboe, flute, and clarinet (using the Kontakt Silver synthesizer in Sibelius 4) as well as purely synthetic stimuli, in order to ascertain the relationship of time-variant spectral properties to consonance, dissonance, and perceived similarity.

2:00

**4pMU3. Multidimensional scaling analysis of centroid- and attack/decay-normalized musical instrument sounds.** James W. Beauchamp (School of Music and Dept. of Elect. & Comput. Eng., Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801, jwbeauch@uiuc.edu), Andrew B. Horner (Hong Kong Univ. of Sci. & Technol., Kowloon, Hong Kong), Hans-Friedrich Koehn, and Mert Bay (Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801)

Ten sustained musical instrument tones (bassoon, cello, clarinet, flute, horn, oboe, recorder, alto saxophone, trumpet, and violin) were spectrally analyzed and then equalized for duration, attack and decay time, fundamental frequency, number of harmonics, average spectral centroid, and presentation loudness. The tones were resynthesized both with time-varying harmonic amplitudes and frequencies (dynamic case) and fixed amplitudes and frequencies (static case). Tone triads were presented to ten musically experienced listeners whose tasks were to specify the most dissimilar and most similar pairs in each triad. Based on the resulting dissimilarity matrix, multidimensional scaling (MDS) was used to position the instruments in two- and three-dimensional metric spaces. Two measures of instrument amplitude spectra were found to correlate strongly with MDS dimensions. For both the static- and dynamic-case 2-D solutions, the ratio of even-to-odd rms amplitudes correlated strongly with one of the dimensions. For the dynamic case, spectral centroid variation correlated strongly with the second dimension. Also, 2-D solution instrument groupings agreed well with groupings based on coefficients of the first two components of a principle components analysis representing 90% of the instruments' spectral variance. [This work was supported by the Hong Kong Research Grants Council's CERG Project 613505.]

2:20

**4pMU4. Sound synthesis based on a new micro timbre notation.** Naotoshi Osaka (School of Eng., Tokyo Denki Univ., 2-2, Kanda-Nishikicho, Chiyoda-ku, Tokyo, 101-8457, Japan, osaka@im.dendai.ac.jp), Takayuki Baba, Nobuhiko Kitawaki, and Takeshi Yamada (Univ. of Tsukuba, Japan)

Timbre has become a major musical factor in contemporary and computer music. However, sufficient timbre theory has not yet been established. The author is challenging to create new timbre theory for music composition. The first step of its construction is to make the timbre descriptive. A micro timbre is defined, which is a perceptual impression of a sound with approximately 50 to 100-ms duration, and describe sound as a micro timbre sequence. This can be used as a new notation system in place of common music notation. In dictation, micro timbre sequence and correspondent duration sequence are perceptually recorded. When synthesizing from this notation, sounds corresponding to the notation systems are either physically synthesized or searched for in a large sound database to generate sound data for a given duration. Two sequential sound data instances are first represented in sinusoidal representations and then are concatenated using a morphing technique. Sounds generated by a stream of water and similar sounds are described using the method as examples. Then scripts describing electronic sounds are introduced and explained. The ability to record, transmit to others, and resynthesize timbre is one of the useful functions of the theory.

2:40

**4pMU5. Timbre representation for automatic classification of musical instruments.** Bozena Kostek (Gdansk Univ. of Technol., Narutowicza 11/12, PL-80-952 Gdansk, Poland)

Human communication includes the capability of recognition. This is particularly true of auditory communication. Music information retrieval (MIR) turns out to be particularly challenging, since many problems remain still unsolved. Topics that should be included within the scope of MIR are automatic classification of musical instruments/phrases/styles, music representation and indexing, estimating musical similarity using both perceptual and musicological criteria, recognizing music using audio and/or semantic description, language modeling for music, auditory scene analysis, and others. Many features of music content description are based on perceptual phenomena and cognition. However, it can easily be observed that most of the low-level descriptors used, for example, in musical instrument classification are more data- than human-oriented. This is because the idea behind these features is to have data

defined and linked in such a way as to be able to use it for more effective automatic discovery, integration, and reuse in various applications. The ambitious task is, however, to provide seamless meaning to low- and high-level descriptors such as timbre descriptors and linking them together. In such a way data can be processed and shared by both systems and people. This paper presents a study related to timbre representation of musical instrument sounds.

### 3:00–3:15 Break

3:15

**4pMU6. An attempt to construct a quantitative scale of musical brightness for short melodies implementing timbral brightness.** Masashi Yamada (Kanazawa Inst. of Technol., 3-1 Yatsukaho, Hakusan, Ishikawa 924-0838, Japan, m-yamada@neptune.kanazawa-it.ac.jp)

It is known that a major tune is brighter than a minor one, and that music played in a faster tempo and a higher register is brighter than a slower and lower one. However, it has not been clarified how these factors quantitatively determine the musical brightness. On the other hand, it has been clarified that the timbral brightness of a tone corresponds well to the spectral centroid. In the present study, major and minor scales and two short melodies were played with pure tones, and listeners evaluated their musical brightness. For each performance, the spectral centroid was calculated for the overall-term spectrum during the performance on the transformed frequency scale of the ERB rate. The results showed that the musical brightness of the ascending scale increases proportionally as the spectral centroid shown in the ERB rate increases. Using this, a quantitative scale of musical brightness, BM, was constructed. The results also showed that the difference in the musical brightness between major and minor scales corresponded to the transposition of approximately 5 ERB rate, and doubling the speed corresponded to the upper shift of the centroid in approximately 2.5 ERB rate.

3:35

**4pMU7. Subjective congruency between a sound effect and a switching pattern of a visual image.** Shinichiro Iwamiya, Motonori Arita, and Sun Su (Dept. of Acoust. Design, Kyushu Univ., 4-9-1, Shiobaru, Minami-ku, Fukuoka 185-8540, Japan)

The relationship between the transformation of a visual image and the pitch pattern of sound can create formal congruency between sounds and moving pictures. This effect by switching patterns of enlarging and reducing images that were combined with ascending and descending pitch scales was examined. Rating experiments showed two congruent patterns of the combination of switching and scale patterns; one was a combination of an ascending pitch scale and an enlarging image pattern, and the other a combination of a descending pitch scale and a reducing image pattern. These forms of matching might be based on a Doppler illusion. An additional pair of congruent patterns for combinations of switching and scale patterns was also found: one was a combination of an ascending pitch scale and a sliding movement from left to right, and the other a combination of a descending pitch scale and a sliding movement from right to left. These forms of matching might be based on the correspondence of a progressive sensation. Further, the formal congruency between a pitch pattern and the formal transportation can contribute to integrating auditory and visual information and to making audio-visual products more impressive.

## Contributed Papers

3:55

**4pMU8. SRA: An online tool for spectral and roughness analysis of sound signals.** Pantelis Vassilakis (School of Music, ITD, Libraries, DePaul Univ., 2350 N. Kenmore Ave., JTR 207, Chicago, IL 60614)

SRA performs spectral and roughness analysis on user-submitted 250- to 1000-ms-long portions of sound files (.wav/.aif formats). Spectral analysis incorporates an improved STFT algorithm [K. Fitz and L. Haken, *J. Aud. Eng. Soc.* **50**(11), 879–893 (2002)] and automates spectral peak-picking using the Loris open source C++ class library [Fitz and Haken (CERL Sound Group)]. Users can manipulate three spectral analysis/peak-picking parameters: analysis bandwidth, spectral-amplitude normalization, and spectral-amplitude threshold. Instructions describe the parameters in detail and suggest settings appropriate to the submitted files and questions of interest. The spectral values obtained from the analysis enter a roughness estimation model [P. N. Vassilakis, *Sele. Rep. in Ethnomusicol.* **12**, 119–144 (2005)], outputting roughness values for each individual sine-pair in the file's spectrum and for the entire file. The roughness model quantifies the dependence of roughness on a sine-pair's (a) intensity (combined amplitude of the sines), (b) amplitude fluctuation degree (amplitude difference of the sines), (c) amplitude fluctuation rate (frequency difference of the sines), and (d) register (lower sine frequency). Presentation of the roughness estimation model and the online tool will be followed by a discussion of research studies employing it and an outline of future possible applications. [Work supported by DePaul University and Eastern Washington University. Programmed by K. Fitz.]

4:10

**4pMU9. Further spectral correlations of timbral adjectives used by musicians.** Alastair C. Disley, David M. Howard, and Andrew D. Hunt (Dept. of Electron., Univ. of York, Heslington, York, YO10 5DD, UK)

As part of a project to develop a synthesis interface which nontechnical musicians should find intuitive, the adjectives musicians use to describe timbre have been studied in a large-scale listening test covering the timbre space of Western orchestral instruments. These were refined in previous work by the authors (A. C. Disley *et al.* "Spectral correlations of timbral adjectives used by musicians," *J. Acoust. Soc. Am.* **119**, 3333, (2006)] to a set of ten words which had good common understanding and discrimination between the samples (bright, clear, dull, gentle, harsh, nasal, percussive, ringing, thin, and warm). To help explore potential relationships between these adjectives and spectral features, 20 listeners participated in a further listening experiment, comparing samples in pairs to produce dissimilarity data. Multidimensional scaling produced dimensions which were compared with a large number of spectral and time-domain analyses of the stimuli, suggesting a number of significantly correlated spectral cues with some of the adjectives. These results are compared with previous studies by the authors and others, showing both similarities and differences, suggesting that collective consideration of timbral adjectives is more likely to result in simultaneously applicable theories of acoustic correlates than individual consideration of words.

**Session 4pNSa****Noise and Architectural Acoustics: Soundscapes and Cultural Perception II**

Brigitte Schulte-Fortkamp, Cochair

*Technical Univ. Berlin, Inst. of Technical Acoustics, Secr TA 7, Einsteinufer 25, 10587 Berlin, Germany*

Bennett M. Brooks, Cochair

*Brooks Acoustics Corp., 27 Hartford Turnpike, Vernon, CT 06066***Contributed Papers****1:00**

**4pNSa1. Mapping soundscapes in urban quiet areas.** Gaetano Licitra (ARPAT, Tuscany Regional Agency for Environ. Protection, Via N. Porpora, 22-50144, Firenze, Italy), Gianluca Memoli (Memolix, Environ. Consultants, 56127 Pisa, Italy), Mauro Cerchiai, and Luca Nencini (ARPAT, 56127 Pisa, Italy)

Innovative action plans in noise-polluted environments require the description of the existing soundscape in terms of suitable indicators. The role of these indicators, giving the “fingerprint” of a fixed soundscape, would be not only to measure the improvement in the sound quality after the action taken, but also to guide the designer in the process, providing a reference benchmark. One of the open questions on new indicators is the way they relate to existing ones and to people’s perception. The present work will describe a “Sonic Garden” in Florence, using both the “slope” indicator (constructed from the  $LA_{eq}$  time history and related in previous studies to people’s perception) and classical psychoacoustical parameters (level, spectral structure, and perceived characteristics such as loudness, sharpness, fluctuation, and roughness). The latter parameters will be acquired using a binaural technique.

**1:15**

**4pNSa2. A questionnaire survey of the attitude of Japanese and foreign residents in Japan to sound masking devices for toilets.** Miwako Ueda and Shin-ichiro Iwamiya (Grad. School of Design, Kyushu Univ., Iwamiya Lab. 4-9-1, Shiobaru, Minami-ku, Fukuoka 815-8540 Japan, amaria@white.livedoor.com)

Unique sound masking devices for toilets can be used in women’s restrooms in Japan. Such devices function to produce the sound of flushing water without actual flushing. To mask the sound of bodily functions, women tended to flush the toilet continuously while using it, thereby wasting a large amount of water. In the circumstances, sound masking devices have been introduced to public toilets. We have recently conducted a questionnaire survey to clarify the attitude of people toward such sound masking devices for toilets. The results of the survey showed that many Japanese women know such devices and often use them, that foreign women currently living in Japan also know that such devices exist, and that some Japanese men have heard of such devices but never used them. Many Japanese women are quite embarrassed at the thought that someone else can hear them while they are on the toilet. Many noted the necessity of such devices and required a wide range of setting for toilets in public spaces. However, they are not satisfied with the sound quality of the play-back toilet flush sounds of currently available devices. The above results suggest that the sound quality of such devices should be improved.

**1:30–2:00****Panel Discussion****Session 4pNSb****Noise, Physical Acoustics, and Structural Acoustics and Vibration: Acoustics of Sports**

Joseph Pope, Cochair

*Pope Engineering Company, P.O. Box 590236, Newton, MA 02459-0002*

Kenji Kurakata, Cochair

*AIST, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan***Chair’s Introduction—2:10****Invited Papers****2:15**

**4pNSb1. A review of the vibration and sounds of the crack of the bat and player auditory clues.** Robert Collier (Thayer School of Eng., 8000 Cummings Hall, Hanover, NH 03755)

The purpose of this paper is to review the state-of-the-art in the acoustics of baseball. As is well known, the crack of the bat is an important phenomenon of solid wood bats and metal bats. Each has a very different sound signature. At the 148th meeting of the ASA in 2004, the author and coauthors Ken Kaliski and James Sherwood presented the results of laboratory and field tests, which showed

that the spectral characteristics of radiated sound are dependent on the ball-bat impact location and resultant bat vibration of both solid wood and tubular metal bats. These results will be reviewed together with those of other investigators in the context of player auditory clues and the player's response in game situations.

2:35

**4pNSb2. Measurements of the impact sound of golf clubs and risk of hearing impairment.** Kenji Kurakata (Nat'l. Inst. of Adv. Industrial Sci. and Technol. (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki, 305-8566 Japan, kurakata-k@aist.go.jp)

The ball-impact sounds of golf clubs with metal heads and of a club with a wood head were measured to investigate their different acoustic properties. Hitting was executed using either a swing machine or a human player. Results of these analyses showed that the metal-head clubs generated sounds around 100 dB ( $L_{pA, Fmax}$ ). This level was 5–15 dB higher than that of the wood-head club. The sounds of the metal-head clubs had greater power in the high-frequency region of 4 kHz and above compared to the wood-head club, which particularly increased the overall sound levels. These results suggest that it would be desirable to develop a metal head with pleasant sound qualities, keeping the sound level lower to minimize hearing damage. Some of these measurement data were published in Japanese in a previous paper [K. Kurakata, *J. INCE/J* **26**, 60–63 (2002)].

2:55

**4pNSb3. New underwater sound system for synchronized swimming: The 9th International Swimming Federation Championships.** Takayuki Watanabe, Shinji Kishinaga (YAMAHA Ctr. for Adv. Sound Technologies, 203 Matsunokijima, Iwata, Shizuoka, 438-0192 Japan), Tokuzo Fukamachi (YAMAHA Motor Marine Operations, Arai, Hamana, 438-8501 Japan), and Osamu Maeda (YAMAHA Motor Adv. Technol. Res. Div., Iwata, 438-8501 Japan)

There have been concerns about the differences between underwater sound fields in a temporary fiberglass-reinforced plastic (FRP) pool and in a conventional reinforced concrete (RC) pool. A temporary FRP pool was to be used for competitions at the World Swimming Championships in Fukuoka. We considered three items as key factors for a swimming pool used for synchronized swimming: (1) the sound source itself (output level, fluctuations in frequency characteristics); (2) the effect of materials used in pool construction upon sound source installation conditions; and (3) the effect of the  $m$ th mode low-frequency cutoff in "shallow water." To improve basic problems related to the first factor, we developed a new actuator-driven underwater sound system (YALAS), which can eliminate the effect of installation conditions for underwater speakers in the FRP pool. This new underwater system has now seen practical use in competitions. The report summarizes this new underwater sound system and compares the system with conventional systems in terms of its acoustic characteristics. The system can offer music with sufficient audibility in water. We gained a good reputation with competitors because the system showed superior performance to conventional systems in sound volume and quality, and in uniformity of sound distribution.

3:15

**4pNSb4. Acoustics of the Great Ball Court at Chichen Itza, Mexico.** David Lubman (14301 Middletown Ln., Westminster, CA 92683)

The ball game has played a central role in Mayan religion and culture for 5000 years. Thousands of ball courts have been discovered. The Great Ball Court (GBC) at Chichen Itza is a late development and is architecturally unique. Two remarkable acoustical features were noticed during excavation in the 1920s, but never explained or interpreted. A whispering gallery permits voice communication between temples located about 460 feet (140 m) apart. A profound flutter echo is heard between the two massive parallel walls of the playing field, about 270 ft (82 m) long, 28 ft (8.5 m) high, and 119 ft (36 m) apart. Until recently, most archaeologists dismissed acoustical features at Mayan sites as unintended artifacts. That is now changing. Stimulated by archaeological acoustic studies and reports since 1999, eminent Mayanists Stephen Houston and Karl Taube have reinterpreted certain Mayan glyphs as vibrant sounds and ballcourt echoes, and have famously called for a new archaeology of the senses, especially hearing, sight, and smell [Cambridge Archaeol. J. **10** (2) 261–294 (2000)]. By interpreting architectural, psychoacoustic, and cognitive features of the GBC in the context of ancient Mayan culture, this paper speculates that acoustical effects at the GBC may be original design features.

### Contributed Papers

3:35

**4pNSb5. Sleep disturbance caused by shooting sounds.** Joos Vos (TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands, joos.vos@tno.nl)

In the present study relations between the sound level of shooting sounds and the probability of behaviorally confirmed noise-induced awakening reactions were determined. The sounds were presented by means of loudspeakers in the bedrooms of 30 volunteers. The shooting sounds had been produced by a small and a medium-large firearm, and the stimuli consisted of individual bangs or volleys of ten isolated or partly overlapping impulses. Aircraft sound was included as a reference source. The sounds were presented during a 6-h period that started 75 min after the beginning of the sleeping period. The time period between the various stimuli varied between 12 and 18 min, with a mean of 15 min. To cope with at least a relevant portion of habituation effects, each subject participated in 18 nights to be completed within 4 weeks. Preliminary results are presented both for the awakening reactions described above, and for vari-

ous other dependent variables collected with the help of an actimeter or determined by means of subjective rating scales. [Work supported by the Dutch Ministry of Defense.]

3:50

**4pNSb6. Sound inside a gymnasium.** Sergio Beristain (ESIME, IPN, IMA., P.O. Box 12-1022, Narvarte, 03001, Mexico City, Mexico)

A new gymnasium for a sports club was designed taking acoustic comfort into consideration, in order to accommodate sports practice, sports events with the public, or musical and drama presentations, taking advantage of its large capacity for the public and performers. The floor plan included room enough for a basketball court with public space on one side, where grades for about 200 people will be permanently installed. Walls were treated in a way that is useful for the sports practice (hard surfaces), with hidden absorption material to reduce the usual reverberant

field inside the court, and to allow for sound events with only the addition of a mat (to protect the floor woodwork) and extra grades and the sound reinforcement system.

4:05

**4pNSb7. Occupational and recreational noise exposures at stock car racing circuits.** Chucri A. Kardous, Thais Morata, and Luann E. Van Campen (Natl. Inst. for Occupational Safety Health, 4676 Columbia Pkwy., Cincinnati, OH 45226, ckardous@cdc.gov)

Noise in stock car racing is accepted as a normal occurrence but the exposure levels associated with the sport have not been adequately characterized. Researchers from the National Institute for Occupational Safety and Health (NIOSH) conducted an exploratory assessment of noise exposures to drivers, pit crew, team staff, and spectators at three stock car racing events. Area measurements were made during race preparation, practice, qualification, and competition. Personal dosimetry measurements were conducted on drivers, crew members, infield staff, and spectators. Findings showed time-weighted averages (TWA) that ranged from 94 decibels A-weighted (dBA) for spectators to 114 dBA for car drivers. Peak sound-pressure levels exceeded the maximum allowable limit of 140 decibels (dB) during race competitions. Personal exposure measurements exceeded the NIOSH recommended exposure limit (REL) of 85 dBA as an 8-h TWA in less than a minute for one driver during practice, within 2 min

for pit crew and infield staff, and 7 to 10 min for spectators during the race. Hearing protection use was variable and intermittent among crew, staff, and spectators. Among drivers and crew, there was greater concern for communication performance than for hearing protection.

4:20

**4pNSb8. Sports acoustics: Using sound from resonant shells, vibrating cylinders, strumming shafts, and water impact to evaluate athletic performance.** David G. Browning (Dept. of Phys., Univ. of Rhode Island, Kingston, RI 02881, decibeldb@aol.com) and Peter M. Scheifele (Univ. of Connecticut, Storrs, CT 06269)

The sound from equipment used and/or specific acts during athletic competition, such as hitting a baseball with an aluminum bat, carries beyond the playing field and can provide a nonobtrusive method to evaluate athletic performance—such as where on the bat the ball was hit. Standardized equipment guarantees repeatability, for example, every volleyball resonates at the same frequency. Each major sport can have unique noise interference which in some circumstances can be overwhelming, and the distance from the sound source can vary significantly during a game. Still, it will be shown that useful performance information can be obtained under realistic conditions for at least the following sports: volleyball, softball, baseball, golf, swimming and diving, soccer, and football.

FRIDAY AFTERNOON, 1 DECEMBER 2006

WAIANAE ROOM, 1:30 TO 6:20 P.M.

### Session 4pPA

## Physical Acoustics and Biomedical Ultrasound/Bioresponse to Vibration: Sound Propagation in Inhomogeneous Media II

Takahiko Otani, Cochair

*Doshisha Univ., Lab. of Ultrasonic Electronics, Kyotonabe-shi, Kyoto 610-0321, Japan*

James G. Miller, Cochair

*Washington Univ., Dept. of Physics, 1 Brookings Dr., St. Louis, MO 63130*

### Invited Paper

1:30

**4pPA1. Observables and prediction modeling in the presence of ultra-wideband heterogeneity.** John J. McCoy (The Catholic Univ. of America, Washington, DC 20064)

Underlying virtually all propagation and scattering models is an intuitive understanding; the acoustic field is observable using a device of a sufficiently small size to obtain a sufficiently dense set of discrete measurements. This assumes the field variation cuts off at an inner length scale larger than the device size, assuring that no information is lost to the inherent spatial averaging in any measurement. This understanding is faulty in the presence of environment heterogeneity observed on an extreme range of length scales. The reason is that all physical devices have finite accuracy, which limits their ability to capture variation on scales significantly larger than their size, in the presence of variation on intermediate scales. A more refined understanding of the ability to observe a field requires multiple devices, an unbounded hierarchy in the limit, to obtain multiple dense sets of discrete “observables.” This, then, suggests a different class of prediction models for environments with ultra-wideband heterogeneity, expressed in multiple sets of discrete variables, each set describing field variation in a limited subband. A framework for formulating these prediction models and their application to a scenario for which environment heterogeneity has no inner scale cutoff is presented.

1:50

**4pPA2. Parameters arising from the Burridge-Keller formulation for poroelastic media, especially for granular media and marine sediments.** Allan D. Pierce, William M. Carey, and Paul E. Barbone (Boston Univ., Boston, MA 02215)

It was previously shown [Pierce *et al.* (2006)] that the Burridge-Keller formulation [J. Acoust. Soc. Am. (1981)] rigorously justifies the low-frequency version [(1956a)] of Biot's equations. Implementation involves two microscale problems: (1) incompressible viscous flow driven in a highly irregular space with rigid walls by a uniformly and externally applied apparent pressure distribution and (2) elastostatic deformation of an intricate elastic web caused by the joint influence of a distributed constant body force and by uniform tractions (including an external pressure) on the web's exposed surface. Microscale averages produce the Biot "constants." Theoretical devices of applied mechanics and mathematics yield estimates of these and related parameters. In particular, it is shown that Wood's equation is a reasonable first approximation for the sound speed in sediments in the low-frequency limit. The formulation also yields an estimation for the sound speed in the high-frequency limit, when the viscous boundary layers become thin. The well-known result that the attenuation varies as  $f^{1/2}$  in the high-frequency limit also results without the necessity of Biot's heuristic patching theory. Various heuristic approximations due to Gassmann, to Geertsma and Smit, and to Stoll and Bryant are analytically and numerically assessed.

2:05

**4pPA3. Sound propagation in the mixtures of liquid and solid aggregate; similarities at micro- and nanoscales..** Hasson M. Tavossi (Dept. of Physical and Environ. Sci., Mesa State College, 1100 North Ave., Grand Junction, CO 81504)

Sound propagation phenomena in certain liquid and solid aggregate mixtures, at micrometer scales, in some cases resemble the wave propagation behaviors of materials observed at nanometer and atomic scales. For example, it can be shown that the sound wave dispersion, attenuation, and cutoff-frequency effects depend on the same structural parameters as those observed at nano or atomic levels and are similar at both scales. Therefore, to investigate theoretical models of wave and matter interactions it is more convenient to use, as experimental tools, the readily analyzable models of wave propagation, in mixtures of solid and liquid, constructed at micrometer scales. Theoretical findings on sound propagation in the mixtures of liquid and solid particles at micrometer scales will be discussed. These results show the resemblance to the behavior of acoustic phonons, the lattice thermal vibrations of crystalline structures, at radically different scales. Experimental data on wave dispersion, attenuation, band-pass, and cutoff frequency effects, measured for sound propagation, in inhomogeneous materials consisting of mixtures of solid and liquid will be presented, showing the similarities of wave propagation behaviors at micro- and nanoscales.

2:20

**4pPA4. Nonlinear surface waves in soil.** Evgenia A. Zabolotskaya, Yurii A. Ilinskii, and Mark F. Hamilton (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029)

Nonlinear effects in surface waves propagating in soil are investigated theoretically. Analytic solutions are derived for the second harmonics and difference frequency waves generated by a bifrequency primary wave propagating at moderate amplitude. The soil is modeled as an isotropic solid. As such, its elastic properties are described by five elastic constants, two at second order in the strain energy density (the shear and bulk moduli) and three at third order. Nonlinear propagation of the surface waves is based on a theory developed previously [Zabolotskaya, J. Acoust. Soc. Am. **91**, 2569–2575 (1992)]. Elements of the nonlinearity matrix associated with the interacting spectral components are expressed in terms of the five elastic constants. It was found convenient to express the nonlinearity matrix for soil as a function of a nonlinearity parameter corre-

sponding to  $B/A$  for liquids, particularly for saturated soils exhibiting liquidlike properties. This nonlinearity parameter can vary by several orders of magnitude. For soils with shear wave speeds less than 20% of the compressional wave speeds, the nonlinearity of surface waves is found to be independent of the third-order elastic constants and dependent only on the shear modulus. [Work supported by ONR.]

2:35

**4pPA5. The measurement of the hysteretic nonlinearity parameter of a field soil by the phase shift method: A long-term survey.** Zhiqu Lu (Natl. Ctr. for Physical Acoust., The Univ. of Mississippi, Univ., MS 38677)

Soil properties significantly affect the performance of the acoustic landmine detection. The climate and seasonal changes cause the variations of soil properties and smear landmine signature over time. On the other hand, soil is a complicated granular material that exhibits strong nonlinear acoustic behaviors. To understand the weather and seasonal effects on nonlinear acoustic behaviors of soils, a phase shift method is used to measure the hysteretic nonlinearity parameter of a field soil. The technique is based on measuring the variation of phase difference between two transducers, i.e., the phase shift, induced by changing sound level. The hysteretic nonlinear parameter can be extracted from the measured phase shift as a function of sound level or dynamic strain. In a long-term survey, the nonlinearity parameter, sound speed, and environmental conditions such as temperature, moisture, soil water potential, and rainfall precipitation are measured. It is found that the nonlinearity parameter is much more sensitive than sound speed to the climate change. Soil water potential is the predominant factor that affects the nonlinearity parameter and sound speed of the shallow field soil.

2:50

**4pPA6. Nonlinear acoustic landmine detection: Comparison of soil nonlinearity with soil-interface nonlinearity.** Murray S. Korman, Kathleen E. Pauls, Sean A. Genis (Dept. of Phys., U. S. Naval Acad., Annapolis, MD 21402), and James M. Sabatier (Univ. of Mississippi, Univ., MS 38677)

To model the soil-top plate interface in nonlinear acoustic landmine detection, the soil-plate oscillator was developed [J. Acoust. Soc. Am. **116**, 3354–3369 (2004)]. A Lexan plate (2.39 mm thick, 18.5 cm diameter) is clamped at an inside diameter of 11.8 cm between two metal flanges. Dry sifted masonry sand (2-cm layer) is placed over the plate. Turning curves experiments are performed by driving a loudspeaker (located over the sand) by a swept sinusoid. The acceleration versus frequency is measured near resonance on a swept spectrum analyzer using an accelerometer centered on the surface. The corresponding backbone curve exhibits a linear decrease in resonant frequency  $f$  versus increasing acceleration, where  $a = -a_0(f - f_0)/f_0$ . Define a nonlinearity parameter  $\alpha = 1/a_0$ . When the elastic plate is replaced by a "rigid" plate,  $\alpha$  decreased from 0.128 to 0.070 ( $s^2/m$ ), while  $f_0$  increased from 191 to 466 Hz. When a cylindrical drum-like mine simulant (rigid walls, thin acrylic top-plate) was buried 2 cm deep in a concrete sand box, "on the mine" results yielded  $\alpha = 0.30$  ( $s^2/m$ ) with  $f_0 = 147$  Hz, while "off the mine,"  $\alpha \sim 0.03$  ( $s^2/m$ ) at  $f_0 = 147$  Hz. [Work supported by ONR.]

3:05

**4pPA7. Causality conditions and signal propagation in bubbly water.** Gregory J. Orris, Dalcio K. Dacol, and Michael Nicholas (Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375)

Acoustic propagation through subsurface bubble clouds in the ocean can exhibit signal travel times with enormous variations depending on the acoustic signal frequency, bubble size distribution, and void fraction. Recent theories have predicted large variations in phase speeds and attenuation that have been largely validated for frequencies well below and well above bubble resonance. However, great care must be exercised when

theoretically treating signal propagation at frequencies near resonance, termed the “Anomalous Absorption Regime” nearly 100 years ago in the pioneering work of Sommerfeld [A. Sommerfeld, *Physik. Z.* **8**, 841 (1975)] while investigating aspects of electromagnetic causality. We will discuss similarities between acoustic propagation in bubbly media and electromagnetic propagation in the presence of a conducting medium. We show that the signal travel time is dependent on the behavior of the dispersion formula in the complex frequency plane and place limits on the range of validity of these formulas, leading naturally to the necessary modifications to the current dispersion formulas to bring them into compliance with causality. Finally, we present theoretical results for the velocity of signals for a representative environment of experimental work carried out at the Naval Research Laboratory. [Work supported by the ONR.]

3:20

**4pPA8. Measurements of the attenuation and sound speed in bubbly salt water.** Gregory J. Orris, Dalcio K. Dacol, and Michael Nicholas (Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375)

Bubble clouds were injected from below the surface of a 144-cubic-meter water tank, wherein hydrophones were placed at varying distances from an acoustic source. Measurements were made over a wide range of frequencies to verify and validate the theoretical predictions of the relevant dispersion formula. This work was undertaken under a variety of conditions by varying the relevant environmental parameters: void fraction, temperature, and salinity. Void fractions were varied from roughly 0.02% to 0.1%. Temperatures ranged from 9 °C to 18 °C, and the salinity was varied from zero to approximately 10% of typical oceanic values. Particular attention was paid to tracking the phase of the transmitted signal as the frequency progressed toward resonance starting from 100 kHz. This yielded phase-speed measurements in an essentially free-field environment using a modified version of phase spectral analysis. Time-of-flight measurements gave signal velocities, while the received energy yielded the attenuation. Results are compared to theoretical calculations, leading to the conclusion that current theoretical dispersion formula requires modification. [This work supported by the ONR.]

3:35

**4pPA9. Efficient computation of 3-D acoustical scattering from multiple arbitrarily shaped objects using the boundary element method/fast multipole method (BEM/FMM).** Nail A. Gumerov and Ramani Duraiswami (Perceptual Interfaces and Reality Lab., Inst. for Adv. Comput. Studies, Univ. of Maryland, College Park, MD 20742)

Many applications require computation of acoustic fields in systems consisting of a large number of scatterers, which may have complex shape. Despite the boundary element method being a well-known technique for solution of the boundary value problems for the Helmholtz equation, its capabilities are usually limited by the memory and speed of computers, and conventional methods can be applicable to relatively small (up to order of 10 000 boundary elements) problems. We developed and implemented an efficient computational technique, based on an iterative solver employing generalized minimal residual method in combination with matrix-vector multiplication speeded up with the fast multipole method. We demonstrate that this technique has  $O(N)$  memory and computational complexity and enables solution of problems with thousands of scatterers (millions of boundary elements) on a desktop PC. The test problems solved are of moderate frequency (up to  $kD \leq 150$ , where  $k$  is the wavenumber and  $D$  is the size of the computational domain). Solution of large scale scattering problems was tested by comparison with the FMM-based T-matrix method applicable for simple shape objects reported earlier [Gumerov and Duraiswami, *J. Acoust. Soc. Am.*, **117**(4), 1744–1761 (2005)], visualization, and physical interpretation of the results.

3:50

**4pPA10. Fast acoustic integral-equation solver for complex inhomogeneous media.** Elizabeth Bleszynski, Marek Bleszynski, and Thomas Jaroszewicz (Monopole Res., 739 Calle Sequoia, Thousand Oaks, CA 91360, elizabeth@monopoleresearch.com)

We describe elements and representative applications of an integral-equation solver for large-scale computations in acoustic wave propagation problems. In the solver construction we used elements of our previously developed fast integral-equation solver for Maxwell’s equations. In comparison with the conventional integral equation approach (method of moments), our solver achieves significant reduction of execution time and memory through the FFT-based matrix compression. One particular aspect of the solver we discuss, pertinent to its high efficiency and accuracy, is an efficient treatment of problems associated with subwavelength discretization. We illustrate the approach and its application on the example of a numerical simulation of acoustic wave propagation through the human head. [Work was supported by a grant from AFOSR.]

4:05

**4pPA11. Models for acoustic scattering in high contrast media.** Max Denis, Charles Thompson, and Kavitha Chandra (Univ. Massachusetts Lowell, One University Ave., Lowell, MA 01854)

In this work a numerical method for evaluating backscatter from a three-dimensional medium having high acoustic contrast is presented. The solution is sought in terms of a perturbation expansion in the contrast amplitude. It is shown that limitations of the regular perturbation expansion can be overcome by recasting the perturbation sequence as a rational fraction using Padé approximants. The resulting solution allows for an accurate representation of the pressure and allows for the poles in the frequency response to be modeled. The determination of the pulse-echo response for a high-contrast medium is discussed and presented.

4:20

**4pPA12. Multiple scattering and visco-thermal effects.** Aroune Duclos, Denis Lafarge, Vincent Pagneux (Laboratoire d’Acoustique de l’Université du Maine, Ave. Olivier Messiaen, 72085 Le Mans, France), and Andrea Cortis (Lawrence Berkeley Natl. Lab., Berkeley, CA 94720)

For modeling sound propagation in a rigid-framed fluid-saturated porous material it is customary to use frequency-dependent density and compressibility functions. These functions, which describe “temporal” dispersion effects due to inertial/viscous and thermal effects, can be computed by FEM in simple geometries and give complete information about the long-wavelength properties of the medium. When the wavelength is reduced, new effects due to scattering must be considered. To study this, we consider solving the sound propagation problem in a 2-D “phononic crystal” made of an infinite square lattice of solid cylinders embedded in a fluid. An exact multiple-scattering solution is first developed for an ideal saturating fluid and then generalized to the case of visco-thermal fluid, by using the concept of visco-thermal admittances. The condition to use this concept is that the viscous and thermal penetration depths are small compared to the cylinder radius. We validate our results in the long-wavelength regime by direct comparisons with FEM data [A. Cortis, “Dynamic parameters of porous media,” Ph.D. dissertation (Delft U.P., Delft, (2002)]. When frequency increases, differences appear between the long-wavelength solution and the exact multiple-scattering solution, which could be interpreted in terms of “spatial” dispersion effects.

4:35

**4pPA13. Effective parameters of periodic and random distributions of rigid cylinders in air.** Daniel Torrent and José Sánchez-Dehesa (Wave Phenomena Group, Nanophotonics Technol. Ctr., Polytechnic Univ. of Valencia, C/Camino de vera s/n., E-46022 Valencia, Spain)

The scattering of sound by finite-size clusters consisting of two-dimensional distributions (periodic and random) of rigid cylinders in air is theoretically studied in the low-frequency limit (homogenization). Analyti-

cal expressions for the effective density and sound speed obtained in the framework of multiple scattering will be reported. For the case of circular-shaped cluster, we have theoretically analyzed the homogenization as a function of the filling fraction, the type of arrangement of the cylinders in the cluster (hexagonal and square lattice), and the number of cylinders in the cluster. When the number of cylinders in the cluster is small we found that for certain “magic numbers” their effective parameters (sound speed and density) are the same as those of the corresponding infinite array. [Work supported by MEC of Spain.]

4:50

**4pPA14. The application of  $k$ -space acoustic propagation models to biomedical photoacoustics.** Benjamin T. Cox (Dept. of Med. Phys. and Bioengineering, Univ. College London, Gower St., London, WC1E 6BT, UK, bencox@mpb.ucl.ac.uk), Simon. R. Arridge, and Paul C. Beard (Univ. College London, Gower St., London, WC1E 6BT, UK)

$k$ -space models for broadband acoustic pulse propagation differ from pseudo-spectral time domain (PSTD) models in their treatment of the time step. By replacing a finite-difference scheme with a propagator, exact for homogeneous media, larger time steps can be taken without loss of accuracy or stability and without introducing dispersion. Three  $k$ -space models for modeling photoacoustically generated (PA) pulses are described here. A very simple, exact, model of PA propagation in a homogeneous fluid is used to introduce the  $k$ -space propagator, and two models of propagation in heterogeneous media, originally designed for modeling scattering in soft tissue, are adapted for use in photoacoustics [Mast *et al.*, IEEE Trans. UFFC **48**, 341–354 (2001); Tabei *et al.*, J. Acoust. Soc. Am. **111**, 53–63 (2002)]. Our motivation for describing these models comes from biomedical PA imaging, in which one of the current limitations is the assumption that soft tissue has a uniform sound speed. Efficient, accurate, and simple-to-encode forward models such as these are very useful for studying the effects of the heterogeneities encountered in practice. They may also be useful in designing PA imaging schemes that can account for acoustic heterogeneities. [This work was funded by the EPSRC, UK]

5:05

**4pPA15. Emergence of the acoustic Green’s function from thermal noise.** Oleg A. Godin (CIRES, Univ. of Colorado and NOAA, Earth System Res. Lab., 325 Broadway, Boulder, CO 80305, oleg.godin@noaa.gov)

Recently proposed applications of noise cross-correlation measurements to passive remote sensing range from ultrasonics and acoustic oceanography to helioseismology and geophysics, at wave frequencies that differ by more than ten orders of magnitude. At the heart of these applications is the possibility to retrieve an estimate of a deterministic Green’s function from long-range correlations of diffuse noise fields. Apparently, S. M. Rytov [A *Theory of Electrical Fluctuations and Thermal Radiation* (USSR Academy of Sciences, Moscow, 1953)] was the first to establish theoretically a simple relation between the Green’s function and the two-point correlation function of fluctuations of wave fields generated by random sources. He used reciprocity considerations to analyze fluctuations of electromagnetic fields. In this paper, an acoustic counterpart of Rytov’s approach is applied to derive exact and asymptotic relations between respective acoustic Green’s functions and cross-correlation of thermal noise in inhomogeneous fluid, solid, and fluid-solid media. Parameters of the media are assumed to be time independent, but can be arbitrary functions of spatial coordinates. Theoretical results obtained are compared to those previously reported in the literature.

5:20

**4pPA16. Simulation of elastic wave scattering in living tissue at the cellular level.** Timothy E. Doyle and Keith H. Warnick (Dept. of Phys., Utah State Univ., 4415 Old Main Hill, Logan, UT 84322-4415, timdoyle@cc.usu.edu)

Elastic wave scattering in biological tissue has been simulated at the cellular level by incorporating a first-order approximation of the cell structure and multiple scattering between cells. The cells were modeled with a concentric spherical shell-core structure embedded in a medium, with the core, shell, and medium representing the cell nucleus, the cell cytoplasm, and the extracellular matrix, respectively. Using vector multipole expansions and boundary conditions, scattering solutions were derived for a single cell with either solid or fluid properties for each of the cell components. Multiple scattering between cells was simulated using addition theorems to translate the multipole fields from cell to cell and using an iterative process to refine the scattering solutions. Backscattering simulations of single cells demonstrated that changes in the nuclear diameter had the greatest effect on the frequency spectra as compared to changes in cell size, density, and shear modulus. Wave field images and spectra from clusters of up to several hundred cells were also simulated, and they exhibited phenomena such as wave field enhancement at the cell membrane and nuclear envelope due to the scattering processes. Relevant applications for these models include ultrasonic tissue characterization and ultrasound-mediated gene transfection and drug delivery.

5:35

**4pPA17. Acoustic analog of electronic Bloch oscillations and Zener tunneling.** José Sánchez-Dehesa, Helios Sanchis-Alepuz, Yu. A. Kosevich, and Daniel Torrent (Wave Phenomena Group, Polytechnic Univ. of Valencia, C/ Camino de Vera s.n., E-46022 Valencia, Spain)

The observation of Bloch oscillations in sound propagation through a multilayer of two different fluidlike components is predicted. In order to obtain the equivalent to the acoustic analog of a Wannier-Stark ladder [E. E. Mendez *et al.*, Phys. Rev. Lett. **60**, 2426–2429 (1988)], a set of cavities with increasing thickness is employed. Bloch oscillations were theoretically predicted as time-resolved oscillations in transmission in direct analogy to electronic Bloch oscillations in semiconductor superlattices [J. Feldmann *et al.*, Phys. Rev. B **46**, R7252–R7255 (1992)]. Finally, an experimental setup is proposed to observe the phenomenon by using arrays of cylindrical rods in air, which acoustically behaves as a fluidlike system with effective sound velocity and density [D. Torrent *et al.*, Phys. Rev. Lett. **96**, 204302 (2006)]. For the proposed system, Bloch oscillations and Zener tunneling are confirmed by using multiple scattering simulations. [Work supported by MEC of Spain.]

5:50

**4pPA18. Comparison of time reversal acoustic and prefiltering methods of focusing of tone burst signals.** Bok Kyoung Choi (Korea Ocean Res. and Development Inst., Sangrok-gu, 426-744, Korea), Alexander Sutin, and Armen Sarvazyan (Artann Labs., Inc., West Trenton, NJ 08618)

The concept of time reversal acoustics (TRA) provides an elegant possibility of both temporal and spatial concentration of acoustic energy in highly inhomogeneous media. TRA-based focusing is typically used for generation of short acoustic pulses, however, in some medical and industrial applications, longer pulses are required. TRA focusing of longer signals leads to an increase of side lobes in temporal and spatial domains. Another method for focusing, known as prefiltering, is based on measurements of the impulse response, which relates the signal at the TRA transmitter to that at the focusing point. After evaluating the impulse response, the excitation signal may be calculated to generate the desired waveform in the focus point. This method allows signal generation with any desired form including long tone-burst signals. Experiments on comparison TRA and prefiltering methods of ultrasound focusing were conducted in the frequency band of 200–1000 kHz. In the experiments, focused acoustic pulses with various forms and duration were generated: triangular, rectan-

gular, and amplitude-modulated tone burst signals. The prefiltering modes provide better temporal compression of the focused signal, and the signal energy outside the main pulse in the prefiltering mode was shown to be much lower than that in standard TRA focusing.

6:05

**4pPA19. Modeling quasi-one-dimensional sound propagation in ducts having two propagation media using a cross-sectional averaging theory.** Donald Bliss and Lisa Burton (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708)

Sound propagation of quasi-one-dimensional waves through a uniform duct partially filled with porous material has been studied theoretically and experimentally. The porous material makes the effective propagation wave number in the duct complex. A fairly simple theory based on cross-

sectional averaging is derived and tested and found to work extremely well up to fairly high frequency. Interestingly, the basic theory depends only on the ratio of cross-sectional areas and the properties of the individual propagation media, but not on the specific configuration of material in a cross section. A higher order correction is developed to achieve excellent accuracy to very high frequency. This correction includes a coefficient that does depend on the specific cross-sectional configuration. Results are compared to exact solutions for layered and annular configurations, and also to experimental measurements with open cell foam as the porous material. An interesting application is to use measured wave numbers to predict the complex effective density and sound speed of porous media samples partially filling the duct. Other applications include fairly simple improved predictions of the behavior of sound in ducts lined with, or partially filled with, bulk reacting absorbing material.

FRIDAY AFTERNOON, 1 DECEMBER 2006

WAIALUA ROOM, 1:30 TO 4:20 P.M.

### Session 4pPP

## Psychological and Physiological Acoustics: Auditory Physiology

G. Christopher Stecker, Cochair

*Univ. of Washington, Dept. of Speech and Hearing Science, 1417 NE 42nd St., Seattle, WA 98105*

Shigeto Furukawa, Cochair

*NTT Communication Science Labs., Human and Information Science Lab., 3-1 Morinosato-wakamiya, Atsugi-shi, Kanagawa-ken 243-0198, Japan*

Chair's Introduction—1:30

### Contributed Papers

1:35

**4pPP1. Transmission of bone-conducted sound measured acoustically and psycho-acoustically.** Sabine Reinfeldt (Signals & Systems, Chalmers Univ. of Technol., SE-412 96 Goteborg, Sweden), Stefan Stenfelt (Linkoping Univ., SE-581 83 Linkoping, Sweden), and Bo Hakansson (Chalmers Univ. of Technol., SE-412 96 Goteborg, Sweden)

The transcranial transmission is important in the bone-conducted (BC) audiometry where the BC hearing thresholds depend on the stimulation position. It is also important for fitting of BC hearing aids; the transcranial transmission determines the amount of the sound that reaches the contralateral cochlea. Previous reported transcranial transmission results seem to depend on the method used. Here, a comparison between the transcranial transmission measured with BC hearing thresholds and ECSP is performed for both open and occluded ear canal. A BC transducer provided stimulation at both mastoids and the forehead. The ECSP was measured with a probe microphone and the BC hearing thresholds were obtained while masking the nontest ear. The transcranial transmission was determined as the BC hearing threshold or the ECSP for contralateral stimulation relative ipsilateral stimulation. The transmission from the forehead was calculated in a similar way. The transcranial transmission was similar for BC hearing thresholds and ECSP above 800 Hz; this indicates that the ECSP can be used as an estimator of the relative hearing perception by BC. The transcranial transmission results are also similar to vibration measurements of the cochleae made in earlier studies. Hence, vibration measurements of the cochleae can also estimate relative BC hearing.

1:50

**4pPP2. Customization of head-related transfer functions using principal components analysis in the time domain.** Ki H. Shin and Youngjin Park (Dept. of Mech. Eng., Korea Adv. Inst. of Sci. and Technol. (KAIST), Sci.e Town, Daejeon, 305-701, Republic of Korea)

Pinna responses were separated from HRIRs (head-related impulse responses) of 45 subjects in the CIPIC HRTF (head-related transfer function) database and modeled as linear combinations of five basic temporal shapes (basis functions) by PCA (principal components analysis) accounting for more than 90% of the variance in the original pinna responses per each selected elevation angle in the median plane. By adjusting the weight of each basis function computed for a specific height to replace the pinna response in the KEMAR HRIR at the same height with the resulting pinna response and listening to the filtered stimuli over headphones, four subjects were able to create a set of median HRIRs that outperformed the KEMAR HRIRs in producing elevation effects in the median plane. Since the monaural spectral features due to the pinna are strongly dependent on elevation instead of azimuth, similar elevation effects could also be generated at different azimuthal positions simply by inserting the customized median pinna responses into the KEMAR HRIRs at other azimuths and varying the ITD (interaural time difference) according to the direction as well as the size of the subject's own head.

2:05

**4pPP3. An electrophysiological measure of basilar membrane nonlinearity in humans.** Christopher J. Plack (Dept. of Psych., Lancaster Univ., Lancaster, LA1 4YF, England) and Ananthanarayan Krishnan (Purdue Univ., West Lafayette, IN 47906)

A behavioral measure of the basilar membrane response can be obtained by comparing the growth in forward masking for maskers at, and well below, the signal frequency. Since the off-frequency masker is assumed to be processed linearly at the signal place, the difference in masking growth with level is thought to reflect the compressive response to the on-frequency masker. The present experiment used an electrophysiological analog of this technique, based on measurements of the latency of wave V of the auditory brainstem response elicited by a 4-kHz, 4-ms pure tone, presented at 65 dB SPL. Responses were obtained in quiet and in the presence of either an on-frequency or an off-frequency (1.8 kHz) pure-tone forward masker. Wave V latency increased with masker level, although the increase was greater for the off-frequency masker than for the on-frequency masker, consistent with a more compressive response to the latter. Response functions generated from the data showed the characteristic shape, with a nearly linear response at lower levels, and 5:1 compression at higher levels. However, the breakpoint between the linear region and the compressive region was at about 60 dB SPL, higher than expected on the basis of previous physiological and psychophysical measures.

2:20

**4pPP4. Possible involvement of the spiral limbus in chinchilla cochlear mechanics.** William S. Rhode (Dept. of Physiol., 1300 University Ave., Madison, WI 53706, rhode@physiology.wisc.edu)

Differences between cochlear mechanical tuning curves and those of auditory nerve fibers (ANFs) exist. In particular, mechanical transfer functions exhibit a high-frequency plateau; ANFs frequency threshold curves (FTCs) do not. ANF-FTCs may have a low-frequency slope due to a velocity forcing function operating on inner hair cells at low frequencies. Neither basilar membrane velocity nor displacement adequately explain the entire ANF tuning curve. A displacement sensitive interferometer was used to study basilar membrane and spiral limbus mechanics in the 6-kHz region of the chinchilla cochlea. The spiral limbus vibrates at the same phase as the basilar membrane nearly up to the location's characteristic frequency. In the plateau region, the limbus appears to vibrate 0 to 20 dB less than the basilar membrane. The basilar membrane/limbus amplitude transfer function has a low-frequency slope of ~3 dB/oct at low frequencies and is ~10 dB lower than the basilar membrane amplitude at 1 kHz. It appears that spiral limbus vibration may contribute to the excitation of the cilia of the inner hair cells. [Work supported by NIDCD grant R01 DC001910.]

2:35

**4pPP5. The effects of antioxidants on cochlear mechanics.** Barbara Acker-Mills, Melinda Hill, and Angeline Ebuon (U.S. Army Aeromedical Res. Lab., P.O. Box 620577, Fort Rucker, AL 36362-0577, barbara.acker@us.army.mil)

Several studies have evaluated the effect of N-acetylcysteine (NAC) on temporary threshold shifts (TTSs) in humans. Work at USAARL found that NAC did not reduce TTSs compared to a placebo, but suppressed otoacoustic emissions (OAEs) more than a placebo, indicating that NAC may reduce outer hair cell activity. Kramer *et al.* [JAAA, 17(4), (2006)] found similar results, where NAC did not affect thresholds in people who had TTSs from exposure to loud music. However, OAEs also did not differ between NAC and placebo. Toppilla *et al.* [XXII Barany Society Meeting (2002)] measured thresholds and balance in people exposed to loud music and found that while NAC did not affect TTS, it reduced noise-induced balance impairment. The current study administered NAC and measured cochlear microphonics, compound action potentials, and OAEs to evaluate cochlear function. The vestibular myogenic potential was measured to assess the effect of NAC on the saccule. The results provide a comprehensive analysis of the effect of NAC on the auditory

system and one component of the vestibular system. [Work supported by the U.S. Army ILIR program. The work is that of the authors and is not necessarily endorsed by the U.S. Army or the Department of Defense.]

2:50–3:05 Break

3:05

**4pPP6. Time characteristics of distortion product otoacoustic emissions recovery function after moderate sound exposure.** Miguel Angel Aranda de Toro, Rodrigo Ordoñez, and Dorte Hammershøi (Dept. of Acoust., Aalborg Univ., Fredrik Bajers Vej 7 B5, DK-9220, Aalborg, Denmark, maat@acoustics.aau.dk)

Exposure to sound of moderate level temporarily attenuates the amplitude of distortion product otoacoustic emissions (DPOAEs). These changes are similar to the changes observed in absolute hearing thresholds after similar sound exposures. To be able to assess changes over time across a broad frequency range, a detailed model of the recovery time characteristics is necessary. In the present study, the methodological aspects needed in order to monitor changes in DPOAEs from human subjects measured with high time resolution are presented. The issues treated are (1) time resolution of the measurements, (2) number of frequency points required, and (3) effects in fine structures, are they affected with the exposure? [Work supported by the Danish Research Council for Technology and Production.]

3:20

**4pPP7. Probability characteristics of neural coincidence detectors in the brainstem.** Ram Krips and Miriam Furst (Dept. of Elec. Eng. Systems, Faculty of Eng., Tel Aviv Univ., Tel Aviv 69978, Israel, mira@eng.tau.ac.il)

Auditory neural activity in the periphery is usually described as non-homogeneous Poisson process (NHPP), characterized as either EE or EI, which operates as a coincidence detector. In order to apply a general probabilistic analysis of those brainstem nuclei activity, the stochastic properties of the axons that exit the EE and EI nuclei is essential. An analytical analysis of the probability characteristics of the output of an EE nucleus (EEout) will be presented. Assuming that an EE nucleus receives inputs from two neurons, each behaves as an NHPP with instantaneous rates  $\lambda_1$  and  $\lambda_2$ , and an output spike is generated if both spike at a coincidence window ( $\Delta$ ) which is relatively small (this matches biological findings). Then EEout is also a NHPP with instantaneous rate  $r(t) = \lambda_1(t) \int_{t-\Delta}^t \lambda_2(\zeta) d\zeta + \lambda_2(t) \int_{t-\Delta}^t \lambda_1(\zeta) d\zeta$ . A similar derivation was applied for an EI nucleus. We found also that the output activity is NHPP for a relatively small coincidence window ( $\Delta$ ). The obtained IR is  $r(t) = \lambda_e(t) [1 - \int_{t-\Delta}^t \lambda_i(\zeta) d\zeta]$ , where  $\lambda_e$  and  $\lambda_i$  are the excitatory and inhibitory input IRs. On the other hand, for larger  $\Delta$ , the output activity is not a Poisson process. Those derivations enable theoretical analysis and performance evaluation of large neural networks, which perform binaural tasks as ITD and ILD.

3:35

**4pPP8. Musical expertise and concurrent sound segregation.** Benjamin Rich Zendel and Claude Alain (Rotman Res. Inst., Baycrest Ctr. & Dept. of Psych., Univ. of Toronto, Toronto, ON M6A 2E1, Canada)

There is growing evidence suggesting that musical training can improve performance in various auditory perceptual tasks. These improvements can be paralleled by changes in scalp recorded auditory evoked potentials (AEPs). The present study examined whether musical training modulates the ability to segregate concurrent auditory objects using behavioral measures and AEPs. Expert musicians and nonmusicians were presented with complex sounds comprised of six harmonics (220, 440, 660 Hz, etc.). The third harmonic was either tuned or mistuned by 1%–16% of its original value. Mistuning a component of a harmonic complex results in the percept of a second auditory object. Stimuli were presented passively (no response) and actively (participants responded by indicating if they heard one sound or two sounds). Behaviorally both musicians and

4p FRI. PM

nonmusicians perceived a second auditory object at similar levels of mistuning. In both groups, complex sounds generated N1 and P2 waves at fronto-central scalp regions. The perception of concurrent auditory objects was paralleled by an increased negativity around 150 ms post-stimulus onset. This increased negativity is referred to as object-related negativity (ORN). Small differences between musicians and nonmusicians were noted in the ORN. The implication of these results will be discussed in terms of current auditory scene analysis theory.

3:50

**4pPP9. Study of auditory temporal processing of language and music in stroke patients.** Li Hsieh and Tamara Baubie (Dept. of Commun. Sci. and Disord., Wayne State Univ., 207 Rackham Hall, 60 Farnsworth, Detroit, MI 48202)

This study focused on functional lateralization of temporal processing of language and music for nine stroke patients and nine normal controls. Subjects were asked to discriminate short versus long sounds in ABX tasks, and then to reproduce sequences of short and long sounds they heard. The reproduction tasks consisted of sequences of 3, 4, and 5 sounds. The linguistic stimuli were nonsense CVC syllables, and the musical stimuli were computer-generated musical notes with the timbres of French horn and trombone. Both linguistic and music stimuli were controlled for frequency and intensity, and varied only for duration (i.e., 500 and 750 ms). Our findings are consistent with previous studies; the left hemisphere specializes in linguistics, while the right hemisphere in music. Moreover, both hemispheres appeared to work closely together in processing temporal information. Both left- and right-hemisphere-damaged patients performed worse than controls. Most subjects performed better with music than language. Patients with left posterior lesions performed worse compared to patients with left or right anterior lesions, particularly with

linguistic stimuli. Patients with right anterior lesions not only involved with temporal processing in music, but also in linguistic stimuli. Our study provided additional information regarding transient temporal processing in language and music.

4:05

**4pPP10. Human bioacoustic biology: Acoustically anomalous vocal patterns used to detect biometric expressions relating to structural integrity and states of health.** Sharry Edwards (Sound Health Res. Inst., P.O. Box 267, Sauber Res. Ctr., Albany, OH 45710)

Computerized analyses of acoustically anomalous vocal patterns are being used as biomarkers for predictive, prediagnostic, and efficient management of individual biological form and function. To date, biometrically distinct vocal data have resulted in outcomes that would be considered improbable by contemporary medical standards. For instance, independent EMG conclusions confirmed the regeneration of nerve tissue for a multiple sclerosis patient who used acoustic bioinformation to guide his primary therapy. Another study monitored the amounts of synthetic labor hormones present during induced labor. False labor costs amount to millions of dollars each year in insurance and hospital resources. The use of noninvasive, possibly remote, vocal profiling could ameliorate such costs. Anomalous vocal acoustics are being investigated by many health-related organizations including Pfizer Pharmaceuticals and the Institute of Automatic Control Engineering in Taiwan. Complementary research studying molecular frequencies of cellular chemistry is being done by James Gimjewski, Ph.D., UCLA, Department of Chemistry and Biochemistry. Known as BioAcoustic Biology, this research modality has the potential to advance current health care standards for biological function, disease processes, and metabolism. Organizations such as the Acoustical Society of America are considering standards for technically defining human bioacoustics. This paper would expand that language.

FRIDAY AFTERNOON, 1 DECEMBER 2006

HONOLULU ROOM, 1:00 TO 3:20 P.M.

### Session 4pSAa

#### Structural Acoustics and Vibration: Vehicle Interior Noise and Vibration

Courtney B. Burroughs, Cochair

*Noise Control Engineering, Inc. 1241 Smithfield St., State College, PA 16801*

Hiroaki Morimura, Cochair

*Japan Science and Technology Agency, Hosei Univ., Dept. of Mechanical Engineering, 3-7-2 Kajino-cho, Kogamei-city, Tokyo, Japan*

#### Invited Papers

1:00

**4pSAa1. Analytical model for flow-excited interior cavity resonance and its application to the Stratospheric Observatory for Infrared Astronomy (SOFIA).** Jerry H. Ginsberg (G. W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

The Stratospheric Observatory for Infrared Astronomy (SOFIA) is a joint effort between NASA and the German Space Agency that has installed a 20 000-kg telescope in a modified 747-SP. The modifications entailed constructing bulkheads, one of which is used to provide the active mount for the telescope, and a door that rotates to open as much as one-quarter of the fuselage circumference to the atmosphere. This configuration represents a large compartment that can exhibit acoustic resonances at low frequencies. Concern arose that a Rossiter mode, which is an aerodynamic resonance in which vortices shed from the leading edge of a gap form a coherent standing pattern at certain speeds, would create a strong acoustic source for acoustic and structural modes, whose frequencies might coincide. A model consisting of a two-dimensional hard-walled waveguide having a Rossiter mode source and an elastic plate at one end was analyzed in order to understand these issues. Unlike conventional analyses of interior cabin noise, in which vibrating walls are the acoustic source, the elastic plate represents a compliant boundary that couples with the acoustic modes. The results lead to some general insights to the possibility of a "triple resonance," as well as the role of structural compliance for cavities that are excited by turbulent external flows.

1:20

**4pSAa2. Energy finite energy analysis for shipboard noise.** Raymond Fischer, Leo Boroditsky (Noise Control Eng. Inc., 799 Middlesex Tnpk, Billerica, MA 01821), Layton Gilfroy (DRDC-Atlantic), and David Brennan (Martec Ltd.)

Machinery-induced habitability noise is difficult to model efficiently and accurately. The potential of energy finite-element analysis (EFEA) is compared to other prediction tools such as statistical energy analysis (SEA). This paper will explore the benefits and costs of EFEA with respect to SEA for acoustic modeling. The focus will be on issues relating to structural modeling for EFEA purposes. EFEA techniques will be evaluated to see if they possess the capabilities of verified SEA approaches for predicting habitability and radiated noise, where it is necessary to account for the impact of diverse marine constructions and sources such as the lack of machinery source information with respect to force or moment inputs or the finite impedance of machinery foundations. The effort proposed herein will provide the necessary engineering to research and identify salient features of EFEA that are potentially applicable for the detailed analysis of the acoustic environment and response of surface ships to various excitation sources. The paper will also address the pros and cons of SEA versus energy-finite element analysis (EFEA) methods used to predict the habitability noise of surface ship platforms. [This work is supported by an Office of Navy Research contract.]

1:40

**4pSAa3. Spectral-based multicoordinate substructuring model for vehicle noise, vibration, and harshness refinement.** Teik C. Lim (Mech., Industrial and Nuclear Eng., 598 Rhodes Hall, Box 210072, Univ. of Cincinnati, Cincinnati, OH 45221, teik.lim@uc.edu)

The success of vehicle NVH (noise, vibration, and harshness) refinement often depends on the ability to identify and understand noise and vibration transmission paths within the mid-frequency range, i.e., 200–1000 Hz, throughout the assembled structure. Due to the complexity of the dynamics in this frequency range, most modal or finite element-based methods do not possess the fidelity needed. To address this gap, a multicoordinate substructuring theory applying measured structural-acoustic and vibration spectra is applied. Three forms of substructuring formulation, namely the nondiagonal, block-diagonal, and purely diagonal coupling cases, are developed. The performances of these approaches are studied numerically, and the net effects of these coupling formulations on the predicted joint and free substructure dynamic characteristics, and system response, are determined. Conditions for applying the simpler coupler that can simplify the testing process and overcome computational deficiencies are also derived. When the measured data is noise contaminated, the singular value decomposition (SVD) algorithm is found to be quite helpful. Using an actual vehicle, a comprehensive analysis of the measured and predicted vehicle system responses is performed. The results are employed to develop an understanding of the primary controlling factors and transfer paths and to cascade system requirements to the substructure level.

2:00

**4pSAa4. Practical application of digital simulation with physical test in vehicle virtual.** Toshiro Abe (ESTECH Corp., 89-1 Yamashita-cho, Naka-ku, Yokohama-shi, Kanagawa-ken, Japan 231-0023, toshiro.abe@estech.co.jp)

In the current vehicle design and development program, the Virtual Product Development process (hereafter, VPD process) is the innovation for automotive industry, which improves product quality and shortens time to market. In general, valid CAE applications are the key component of the VPD process as well as physical tests being indispensable to create valid simulation technologies. This presentation explains how physical-test-based CAE leverages the VPD process. In particular, the presentation is based on how physical testing supports the VPD process and why the ESTECH philosophy is that “The essence of CAE lies in its synergy with Testing;” a philosophy that differentiates the company from the competition. To demonstrate these activities, case studies based on automotive dynamic and real time simulation will be presented. In the case studies, vehicle body NVH and brake noise analysis will be used to show the interaction between physical test and computer simulation. Finally, practical application of the VPD process in one of the leading Japanese automotive companies will be shown, where the effectiveness of the front loading in the actual vehicle development program and the actual deployment of VPD process to the Functional Digital Vehicle project in the powertrain design are to be presented.

### Contributed Papers

2:20

**4pSAa5. Active noise control simulations using minimization of energy density in a mock helicopter cabin.** Jared Thomas, Stephan P. Lovstedt, Jonathan Blotter, Scott D. Sommerfeldt (Brigham Young Univ., Provo, UT 84602), Norman S. Serrano, and Allan Egelston (Silver State Helicopters, Provo, UT 84601)

Helicopter cabin noise is dominated by low-frequency tonal noise, making it an ideal candidate for active noise control. Previous work in active control of cabin noise suggests an energy density approach to be a good solution [B. Faber and S.D. Sommerfeldt, Global Control in a Mock Tractor Cabin Using Energy Density, Proc. ACTIVE 04, Sept. 2004.] Simulations for active noise control using energy density minimization have been made using recorded data from a Robinson R44 helicopter. Initial computer models show substantial noise reductions in excess of 6 dB at the error sensor are possible. Performance results for computer models and simulations in a mock cab for different control arrangements will be compared.

2:35

**4pSAa6. Modeling airborne interior noise in full vehicles using statistical energy analysis.** Arnaud Charpentier and Phil Shorter (ESI, 12555 High Bluff Dr., Ste. 250, San Diego, CA 92130, arnaud.charpentier@esi-group-na.com)

SEA is particularly well suited for predicting airborne noise in vehicles. The acoustic sources found in such environment are typically spatially distributed around the vehicle and can be well represented with SEA diffuse acoustic loads. Multiple transmission paths contribute to interior noise levels including (1) mass law transmission through trimmed panels, (2) resonant radiation from vibrating structures, and (3) flanking paths through gaskets, seals, and holes. All these transmission mechanisms may be modeled using SEA techniques. Finally, interior trim (including carpet, headliner, seats) is a key contributor to the acoustic performance of modern vehicles. The vehicle sound package has a significant impact on both the strength of the transmissions paths into the vehicle as well as the acoustic absorption in the cabin. Both these effects can be accounted for with SEA through detailed models of the trim. SEA models of full vehicles

4p FRI. PM

are usually validated against experimental results at both component and system levels. The models can then be confidently used to (a) rank key design parameters governing interior levels and (b) quickly evaluate the impact of potential design changes. Example vehicle models and correlation results are presented here.

2:50

**4pSAa7. Radiation from vibrating panels at high frequency including an inquiry into the role of edges and drive points.** Donald Bliss (Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708)

In the high frequency limit, a vibrating finite panel is shown to have broadband power and directivity characteristics that can be expressed analytically by a limited set of parameters. Two-dimensional problems with subsonic structural wave speed are considered. Three basic directivity patterns are identified, associated with right and left traveling waves and the correlation between them. The role of boundary conditions at the panel edge is illustrated, as are the effects of types of forcing. Overall, relatively simple broadband behaviors are revealed. The analytical characterization of the radiation is shown to be particularly straightforward in the high frequency broadband limit. Interestingly, the radiated mean-square pressures are independent of the panel length, indicating the radiation is associated with the edges and the drive point. However, the radiation patterns cannot be explained in terms of simple volumetric sources placed just at the edges and the drive point, showing that the often-stated idea of uncan-

celed volumetric sources at these locations is not correct except under very restricted circumstances. A correct physical interpretation of the radiation is provided both in physical space and in terms of spatial Fourier transforms.

3:05

**4pSAa8. Some effect of trim and body panels on the low-frequency interior noise in vehicles.** Andrzej Pietrzyk (Volvo Car Corp., Noise and Vib. Ctr., 405 31 Gothenburg, Sweden)

Structure borne noise dominates the interior noise in vehicles at low frequencies. One of the basic vibroacoustic characteristics of the trimmed body is the noise transfer function, i.e., the acoustic response at a selected position in the passenger compartment, e.g., driver's ear due to a mechanical excitation at a selected body mount. Detailed CAE models based on the FE method are today available for calculating this characteristic at low frequencies, corresponding to the engine idling and road excitation. However, the accuracy of CAE predictions of interior noise is still considered insufficient for the so-called analytical sign off, i.e., zero-prototypes vision. The current paper describes some investigations into the contribution of individual body panels to the overall interior noise. Also, the effect of selected interior trim items on the area investigated. Relative errors of prediction at different trim levels and different frequency ranges are discussed. Both experimental and CAE results are provided. The aim is to better understand the way the interior noise in vehicles is created, and how it can be controlled.

FRIDAY AFTERNOON, 1 DECEMBER 2006

HONOLULU ROOM, 3:45 TO 5:45 P.M.

### Session 4pSAb

## Structural Acoustics and Vibration: General Vibration and Measurement Technology

Peter C. Herdic, Cochair

*Naval Research Lab., Physical Acoustics Branch, Code 7136, 4555 Overlook Ave., SW, Washington, DC 20375*

Naoya Kojima, Cochair

*Yamaguchi Univ., Dept. of Mechanical Engineering, 2-16-1 Tokiwadai, Ube, Yamaguchi 755-8611, Japan*

### Contributed Papers

3:45

**4pSAb1. Direct computation of degenerate elastodynamic solution of elastic wave propagation in a thick plate.** Jamal Ghorieshi (Eng. Dept., Wilkes Univ., Wilkes-Barre, PA 18766)

The limiting form of elastodynamic solutions as frequency tends to zero leads to the elastostatic eigenvalue equations. However, this limiting procedure is not convenient. It is cumbersome when applied to the solutions obtained using Stokes potentials and, in the case of utilizing Lame potentials, it does not produce static solutions that are a function of position alone. In this paper it is shown that the exact solutions of elastostatic problems can, in general, be obtained in a straightforward manner by the use of harmonic potentials without recourse to any special limiting form of analysis. This method is applied to an infinite, elastic thick plate with traction-free parallel surfaces and the elastostatic eigenvalue equation. It is shown that the problem can be solved exactly in terms of harmonic functions, one of which is a scalar and the other one is a vector. It is noted that results are in agreement with the published solutions.

4:00

**4pSAb2. Wave propagation characteristics of an infinite fluid-loaded periodic plate.** Abhijit Sarkar and Venkata R. Sonti (Indian Inst. of Sci. Mech. Eng., IISc., Bangalore-560012, India)

A 1-D infinite periodic plate with simple supports placed along equidistant parallel lines is considered using the finite-element method. The plate is loaded with a finite-height fluid column covered on the top with a rigid plate. Results show a relation between the propagation constant of the fluid-loaded structure with its *in vacuo* counterpart. Since the acoustic medium is an additional wave carrier, the attenuation bands corresponding to the *in vacuo* structure turn out to be propagating. However, the presence of the fluid can also bring about attenuation regions within the *in vacuo* propagation bands. Primary propagation constants bring additional waves called space harmonics with them. Hence, a localized coincidence effect is seen where a particular harmonic falls below or above the acoustic wave number, leading to propagation or a mass loading effect. Occasionally, a complete attenuation band is created. This is verified by decomposing the

single span displacement profile into the space harmonics and also by computing the frequency response function (FRF) for a finite fluid-loaded periodic plate and observing the huge antiresonance dip in frequency in the exact same frequency band where an attenuation band was predicted for the infinite structure.

4:15

**4pSab3. The dynamic response of a plate subjected to various edge excitations.** Baruch Karp (Faculty of Aeronaut. Eng., Technion Israel Inst. of Technol., Haifa 32000, Israel)

Plane strain response of a semi-infinite, elastic plate to harmonic edge excitation is investigated analytically. The exact solution to this problem is obtained as a series expansion of the Rayleigh-Lamb modes of the plate. The variation of energy partition among the propagating modes with frequency of the edge excitation was found for load and displacement (symmetrical) perturbations of uniform and cosine form. The biorthogonality relation was employed in deriving the relative amplitudes of each mode to the given perturbation. The emphasis here is on the sensitivity of the far-field response, represented by propagating waves, to the details of the excitation. Within the frequency range investigated it was found that the plate's response is remarkably insensitive to whether the excitation is load or displacement type. The two types of edge excitation distributions, on the other hand, result in different patterns of energy partition above the first cutoff frequency with similar energy partition only within limited range of frequencies. The effect of the nature of the excitation on the dynamic response of the plate and a possible implication to dynamic equivalence is discussed.

4:30

**4pSab4. Measurement of structural intensity using patch near-field acoustical holography.** Kenji Saijyou (Fifth Res. Ctr., TRDI, Japan Defense Agency, 3-13-1 Nagase, Yokosuka City, Kanagawa Prefecture, Japan, 239-0826)

Measurement of power flow in a structure, called the structural intensity (SI), is essential for vibration control and noise reduction. The near-field acoustical holography (NAH)-based measurement method is suitable to analyze the interrelationship between SI and acoustic intensity (AI) because NAH-based methods provide measurements of SI and AI simultaneously. Use of NAH requires the measurement of a pressure field over a complete surface located exterior to the structure. Therefore, if the measurement aperture is smaller than the structure, reconstructed results from the pressure on the finite aperture are seriously contaminated. This finite aperture effect prevents implementation of this NAH-based method on an actual large-scale structure such as a ship. Patch NAH and regularization method for SI measurement are applied to overcome this difficulty. This method enables implementation of the NAH-based SI measurement method in a large-scale structure. The effectiveness of this method is demonstrated by experiment.

4:45

**4pSab5. Flexural component and extensional component of vibration energy in shell structure.** Taito Ogushi, Manabu Yahara, Masato Mikami, and Naoya Kojima (Dept. of Mech. Eng., Yamaguchi Univ., 2-16-1 Tokiwadai, Ube, Yamaguchi 755-8611, Japan, j008ve@yamaguchi-u.ac.jp)

In this research, the behavior of the flexural component and the extensional component of vibration intensity and their transmission in curved shells are presented. L-shaped shell model was employed as an analysis model of FEM. As FEM analysis methods, both the frequency response

analysis and the transitional response analysis were employed. The flexural component and the extensional component of vibration intensity (VI) were calculated by the results of FEM analysis. In the flexural component of the VI, the vibration energy supplied in the flat part decreased at the boundary from the flat part to the curved part and VI vectors flew in circumferential direction in the curved part. In the extensional component of the VI, the vibration energy appeared at the boundary from the flat part to the curved part and most VI vectors flew parallel to the shell axis in the curved part. The total vibration energy of the flexural component and the extensional component was conserved. So, the vibration energy transformed to each other between the flexural component and the extensional component in L-shaped shell.

5:00

**4pSab6. Seismic/acoustic detection of ground and air traffic for unattended ground sensor technology.** Peter C. Herdic (Naval Res. Lab., Physical Acoust. Branch, Washington, DC 20375 and SFA Inc., Crofton, MD), Brian H. Houston, Phillip A. Frank, and Robert M. Baden (Naval Res. Lab., Washington, DC 20375)

Human footfall and vehicle traffic create surface waves in soil media that can easily be detected by seismic sensors. Field measurement data have been acquired with a triaxial geophone at several experimental sites. The in-plane-surface wave components dominate the response and decay at a rate of approximately  $1/R$ , where  $R$  is distance. This decay rate is due to the combined effect of spreading ( $1/\sqrt{R}$ ) and damping losses in the soil. Further, the detection range is dependent upon the level of environmental noise, soil compliance, moisture content, and topography. Human detection was achieved in rural environments at distances up to  $\sim 30$ – $40$  m, and vehicle detection was possible at much greater distances. Seismic signals due to aircraft are small when compared to the acoustic signature. Ground-based microphone measurements clearly show the blade passage frequency tones of propeller airplanes and the broader band signature of turbojet aircraft. Time- and frequency-domain signal-processing methods for the detection and identification will also be introduced. These experimental results will be discussed with particular emphasis placed on wave phenomenon, detection and identification algorithms, and the related physics.

5:15

**4pSab7. Modeling of acoustic and elastic wave phenomena using plane wave basis functions.** Tomi Huttunen, Jari P. Kaipio (Dept. of Phys., Univ. of Kuopio, P.O. Box 1627, FI-70211 Kuopio, Finland), and Peter Monk (Univ. of Delaware, Newark, DE 19716)

When simulating acoustic, elastodynamic, or coupled fluid-solid vibration problems using standard finite element (FE) techniques, several elements per wavelength are needed to obtain a tolerable accuracy for engineering purposes. At high wave numbers, the requirement of dense meshes may lead to an overwhelmingly large computational burden, which significantly limits the feasibility of FE methods for the modeling of wave phenomena. A promising technique for reducing the computational complexity is to use plane wave basis functions opposed to the low-order polynomials that are used in conventional FE methods. A possible method for utilizing the plane wave basis is the ultra-weak variational formulation (UWVF). The UWVF method can be used for acoustic Helmholtz problems, elastodynamic Navier problems, or fluid-solid systems characterized by a coupled Helmholtz-Navier model. A comparison of the UWVF technique with a low-order FE method shows reduced computational complexity and improved accuracy.

Poster paper 4pSAb8 will be on display from 1:00 p.m. to 5:45 p.m. The author will be at the poster from 5:30 p.m. to 5:45 p.m.

**4pSAb8. Application of vibration energy flow to evaluation of thickness.** Akihiko Higashi (Dept. of Maritime Sci. and Technol., Japan Coast Guard Acad., 5-1 Wakaba-cho, Kure, Hiroshima, 737-8512, Japan)

In this study, the possibility of the useful application of the vibration energy flow is investigated. The vibration energy flow means the propagation of the vibration energy of the flexural waves. The vibration energy flow is expressed by the structural intensity. Here, it is easy to input the flexural waves in the thin plates and beam elements. Then, large structures such as ships use many of these thin plates and beam elements. But the

usual methods of the evaluation and the inspection of these large structures are inefficient. Then, we investigated the possibility of the evaluation of the changed thickness of the structure by using the vibration energy flow analysis. As the result of analysis, the structural intensity suddenly changes at the position of the changed thickness. The changed quantity of the structural intensity corresponds to the change quantity of the thickness. Then, the evaluation method of the thickness of the structure is proposed. As a result, it was found that the change of the structural intensity indicates the change of the thickness. And the evaluation of the change of thickness of beams could be possible by using the proposed method.

FRIDAY AFTERNOON, 1 DECEMBER 2006

MOLOKAI ROOM, 1:00 TO 4:00 P.M.

### Session 4pSC

## Speech Communication: Variation in Production and Perception of Speech (Poster Session)

Heriberto Avelino, Cochair

*Univ. of California—Berkeley, Dept. of Linguistics, 1203 Dwinelle Hall, Berkeley, CA 94720-2650*

Haruo Kubozono, Cochair

*Kobe Univ., Dept. of Linguistics, Faculty of Letters, Nada-ku, Kobe 657-8501, Japan*

### Contributed Papers

All posters will be on display from 1:00 p.m. to 4:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:00 p.m. to 2:30 p.m. and contributors of even-numbered papers will be at their posters from 2:30 p.m. to 4:00 p.m.

**4pSC1. Cross-language perception of voice and affect.** Christer Gobl, Irena Yanushevskaya, and Ailbhe N. Chasaide (Phonet. and Speech Lab., School of Linguistic, Speech and Commun. Sci., Trinity College Dublin, Dublin 2, Ireland, yanushei@tcd.ie)

The paper reports on a cross-language study of how voice quality and  $f_0$  combine in the signaling of affect. Speakers of Irish-English and Japanese participated in perception tests. The stimuli consisted of a short utterance where  $f_0$  and voice source parameters were varied using the LF-model implementation of the KLSyn88a formant synthesizer, and were of three types: (1) VQ only involving voice quality variations and a neutral  $f_0$  contour; (2)  $f_0$  only, with different affect-related  $f_0$  contours and modal voice; (3) VQ+ $f_0$  stimuli, where the voice qualities of (1) combine with specific  $f_0$  contours from (2). Overall, stimuli involving voice quality variation were consistently associated with affect. In (2) only stimuli with high  $f_0$  yielded high affective ratings. Striking differences emerge between the ratings obtained from the two language groups. The results show that not only were some affects consistently perceived by one language group and not the other, but also that specific voice qualities and pitch contours were associated with very different affects across the two groups. The results have important implications for expressive speech synthesis, indicating that language/culture-specific differences need to be considered. [This work is supported by the EU-funded Network of Excellence on Emotion, HUMAINE.]

**4pSC2. An articulatory study of coronal consonants in Arrernte.** Marija Tabain (LaTrobe Univ., Melbourne, Australia), Richard Beare (Monash Univ., Melbourne, Australia), Catherine Best (Univ. of Western Sydney, Sydney, Australia), and Louis Goldstein (Haskins Labs., CT)

This paper presents electro-magnetic articulography (EMA) data on the four coronal stops of Arrernte, an Australian language. The stops are: the lamino-dental “th,” the apico-alveolar “t,” the apico-postalveolar (or “retroflex”) “rt,” and the lamino-palatal “ty.” Jaw, tongue tip (TT), and tongue body (TB) data were collected for two female speakers of the language. Results for the first speaker show a fronted tongue position for the laminal consonants, with the TT reflecting a similar location for both the dental and the palatal. However, for the palatal, the TB position is much higher, whereas for the dental, the TB is very low. For the apical consonants, the TT is not as far forward, and the TB is not quite as high as for the lamino-palatal. For both TT and TB, apico-postalveolar is further back than apico-alveolar. For the second speaker, the TT sensor failed, but in line with the first speaker, the TB sensor showed a higher position for the palatal. The other stops were lower and more forward, with the post-alveolar TB position higher than the laminal or alveolar stop position. For both speakers, the jaw position is lowest for the postalveolar. [Work supported by Australian Research Council and NIH: NIDCD.]

**4pSC3. Symbolic phonetic features for pronunciation modeling.** Rebecca A. Bates,<sup>a)</sup> Mari Ostendorf (Dept. of Elec. Eng., Univ. of Washington, Box 352500, Seattle, WA 98195), and Richard A. Wright (Univ. of Washington, Seattle, WA 98195)

A significant source of variation in spontaneous speech is due to in-traspeaker pronunciation changes, often realized as small feature changes, e.g., nasalized vowels or affricated stops, rather than full phone transformations. Previous computational modeling of pronunciation variation has typically involved transformations from one phone to another, partly because most speech processing systems use phone-based units. Here, a phonetic-feature-based prediction model is presented where phones are represented by a vector of symbolic features that can be on, off, unspecified, or unused. Feature interaction is examined using different groupings of possibly dependent features, and a hierarchical grouping with conditional dependencies led to the best results. Feature-based models are shown to be more efficient than phone-based models, in the sense of requiring fewer parameters to predict variation while giving smaller distance and perplexity values when comparing predictions to the hand-labeled reference. A parsimonious model is better suited to incorporating new conditioning factors, and this work investigates high-level information sources, including both text (syntax, discourse) and prosody cues. Detailed results are under review with *Speech Communication*. [This research was supported in part by the NSF, Award No. IIS-9618926, an Intel Ph.D. Fellowship, and by a faculty improvement grant from Minnesota State University Mankato.] <sup>a)</sup>Currently at Minnesota State University, Mankato.

**4pSC4. Acoustic phonetic variability and auditory word recognition by dyslexic and nondyslexic children.** Patricia Keating, Kuniko Nielsen (Phonet. Lab., Linguist., UCLA, Los Angeles, CA 90095-1543, keating@humnet.ucla.edu), Frank Manis, and Jennifer Bruno (USC, Los Angeles, CA 90089)

The hypothesis that dyslexia involves degraded phonological representations predicts impairments in behaviors that rely on these representations, such as auditory word recognition. Normal adult listeners recognize different pronunciations of a word as instances of the same lexical item, but more slowly and less accurately; dyslexics should be even more impaired by acoustic phonetic variability. Children with and without dyslexia performed a word recognition task: on each trial, a child hears a target word, then eight probes (matching the target or not), responding yes/no to each probe. On some trials, probes are spoken by multiple talkers who differ in age, sex, speech style, etc.; on some trials the match probes also differ from the target in final stop consonant allophone. Responses are scored for accuracy and speed. Research questions include: Do all children demonstrate less accurate/slower recognition of words spoken by multiple talkers versus by one talker? Do all children demonstrate less accurate/slower recognition of words spoken with different allophones? Do dyslexic children demonstrate less accurate/slower recognition than nondyslexic children and, if so, for all trials, only for multiple talker trials, and/or only for different allophone trials; for all dyslexic children, or only those with particular phonological impairments? [Work supported by NIH.]

**4pSC5. Intertalker differences in intelligibility of cochlear-implant simulated speech.** Tessa Bent, Adam B. Buchwald, and David B. Pisoni (Indiana Univ., Dept. of Psychol. and Brain Sci., 1101 E. 10th St., Bloomington, IN 47405, tbent@indiana.edu)

Are the acoustic-phonetic factors that promote highly intelligible speech invariant across different listener populations? Two approaches have been taken to investigate intelligibility variation for a variety of listener populations including hearing-impaired listeners, second language learners, and listeners with cochlear implants: studies on how speaking style affects intelligibility and other research on how inherent differences among talkers influence intelligibility. Taking the second approach, we compared intertalker differences in intelligibility for normal-hearing listeners under cochlear implant (CI) simulation ( $n=120$ ) and in quiet ( $n$

$=200$ ). Stimuli consisted of 20 native English talkers' productions of 100 sentences. These recordings were processed to simulate listening with an eight-channel CI. Both clear and CI-processed tokens were presented to listeners in a sentence transcription task. Results showed that the most intelligible talkers in quiet were not the most intelligible talkers under CI simulation. Furthermore, listeners demonstrated perceptual learning with the CI-simulated speech but showed little learning in the quiet. Some of the acoustic-phonetic properties that were correlated with intelligibility differed between the CI-simulated speech and the speech in the quiet. These results suggest that the intertalker variations that result in highly intelligible speech observed in earlier studies are dependent on listener characteristics. [Work supported by NIDCD.]

**4pSC6. The effect of phonological neighborhood density and word frequency on vowel production and perception in clear speech.** Rajka Smiljanic, Josh Viau, and Ann Bradlow (Dept. of Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208)

Previous research showed that phonological neighborhood density and word frequency influence word recognition (Luce and Pisoni, 1998) and vowel production (Wright, 2002; Munson and Solomon, 2004; Munson, to appear), suggesting an interaction of lexical and phonetic factors in speech production and perception. Here, we explore whether hyperarticulated, intelligibility-enhancing clear speech shows similar sensitivity to lexical-level structure. Nine American English talkers (five females, four males) produced 40 monosyllabic easy (frequent words with few lexical neighbors) and hard (infrequent words with many lexical neighbors) words in conversational and clear speech. Twenty-four subjects participated in a word-in-noise listening test. Results revealed a large effect of style on intelligibility and vowel production: words were more intelligible and vowels were longer and more dispersed in clear compared to conversational speech. Moreover, the female talkers produced larger vowel spaces than male talkers in both speaking styles. Vowels in hard words were marginally more dispersed than vowels in easy words in both speaking styles. However, within both speaking styles, easy and hard words were equally intelligible and of approximately equal duration. These results showed that phonetic properties of vowels were enhanced equally in clear speech regardless of their lexical properties.

**4pSC7. Phoneme dependency of accuracy rates in familiar and unknown speaker identification.** Kanae Amino, Takayuki Arai (Dept. of Elec. and Electron. Eng., Sophia Univ., 7-1 Kioi-cho, Chiyoda-ku, Tokyo, 102-8554 Japan, amino-k@sophia.ac.jp), and Tsutomu Sugawara (Sophia Univ., Chiyoda-ku, Tokyo, 102-8554 Japan)

For perceptual speaker identification, the identification accuracy depends on the speech contents presented to subjects. Our previous studies have shown that stimuli containing nasals are effective for identifying familiar speakers [Amino *et al.*, *Acoust. Sci. Tech.* **27**(4) (2006)]. We have also presented the possibility that the interspeaker spectral distances reflect perceptual speaker similarities. In the present study, we conducted an experiment in which four unknown speakers were identified by 15 subjects. The stimuli were identical to those used in the previous study, in which ten speakers were identified by familiar listeners, although the speakers were fewer this time. Nine consonants in the CV structure were used as stimuli. The consonants were /d/, /t/, /z/, /s/, /r/, /j/, /m/, /n/, and /nj/; the vowel was restricted to /a/ for all CV syllables to simplify the experiment. The results showed that the nasals /n/ and /nj/ obtained higher scores. Tendencies in the differences among consonants were on the same order as those of the

previous experiment, but the average scores were lower than those for familiar listeners. [Work supported by Grant-in-Aid for JSPS Fellows 17-6901.]

**4pSC8. Speech style and stereotypical character in Japanese.** Akiko Nakagawa (Grad. School of Cultural Studies and Human Sci., Kobe Univ., 1-2-1 Tsurukabuto, Nada-ku, Kobe 657-8501, Japan, akiko.nakagawa@atr.jp) and Hiroko Sawada (Kyoto Univ., Kyoto 606-8501, Japan)

This study shows that “stereotypical character” is necessary to understand Japanese speech communication in addition to existing conceptions such as emotion, communicative strategy, register, and so on. Stereotypical character is here defined as a complex entity, consisting of information about gender, age, social status, physical features, characteristics, and speech style. The necessity of stereotypical character was shown through an auditory experiment involving a total of 70 speech sounds comprised of 15–19 short phrases (mean duration 1.4) selected from recordings of spontaneous speech of four adult female speakers of Japanese. Ten participants were asked to listen to these speech sounds randomly, and to classify them into four speakers. Each of the resulting auditory-perceptual categories was found to contain speech sounds from more than one speaker. Further analyses of these results suggested that the participants classified the speech sounds not according to invariant speaker characteristics but according to virtual stereotypical characters that are common in Japanese society. Therefore, such changeable speaker characteristics as “busybody” “thoughtful,” “high-handed,” and so on, can be elicited through speech sounds by Japanese speakers. [This work was partially supported by the Ministry of Education, Science, Sport, and Culture, Grant-in-Aid for Scientific Research (A), 16202006.]

**4pSC9. Perceived vocal age and its acoustic correlates.** Hiroshi Kido (Dept. of Commun. Eng., Tohoku Inst. of Technol., Taihaku-ku, Sendai, Japan 989-8577, kidoh@toitech.ac.jp) and Hideki Kasuya (Intl. Univ. of Health and Welfare, Otawara, Japan 324-8501)

This study investigates relationships between perceived and chronological age of talkers and acoustic correlates of the perceived age. Most of the past studies were primarily concerned with the instability of the vocal-fold vibration extracted from sustained vowels. This study focuses on the dynamic nature of sentence utterances. Talkers included 115 healthy men, aged 20–60 years, who read a short sentence in Japanese. Listeners consisted of 70 men and women, aged 20–40 years, who made direct estimations of age. The results showed a strong correlation ( $r=0.66$ ) between the perceived and chronological age as well as the tendency toward overestimating the ages of younger talkers and underestimating those of older talkers, supporting past investigations [e.g., R. Huntley *et al.*, *J. Voice* **1**, 49–52 (1987)]. Acoustic parameters considered were median of the fundamental frequency ( $F_0$ ) contour,  $F_0$  range, declination of  $F_0$  contour, spectral tilt, median of the boundary frequencies above which irregularities dominate, and speaking rate. From both statistical graphical modeling and regression tree analysis, the speaking rate,  $F_0$  declination, and spectral tilt were found to be dominant acoustic correlates to the perceived age. [Work supported partly by a Grant-in-Aid for Scientific Research, JSPS (16300061).]

**4pSC10. A cross-linguistic study of informational masking: English versus Chinese.** Bruce A. Schneider, Liang Li, Meredyth Daneman (Dept. of Psych., Univ. of Toronto at Mississauga, Mississauga, ON, L5L 1C6 Canada, bschneid@utm.utoronto.ca), Xihong Wu, Zhigang Yang, Jing Chen, and Ying Huang (Peking Univ., Beijing, China 10087)

The amount of release from informational masking in monolingual English (Toronto, Canada), and Chinese (Beijing, China) listeners was measured using the paradigm developed by Freyman *et al.* [*J. Acoust. Soc. Am.* **106**, 3578–3588]. Specifically, psychometric functions relating

percent-correct word recognition to signal-to-noise ratio were determined under two conditions: (1) masker and target perceived as originating from the same position in space; (2) masker and target perceived as originating from different locations. The amount of release from masking due to spatial separation was the same for English and Chinese listeners when the masker was speech-spectrum noise or cross linguistic (Chinese speech masking English target sentences for English listeners or English speech masking Chinese target sentences for Chinese listeners). However, there was a greater release from masking for same-language masking of English (English speech masking English target sentences) than for same-language masking of Chinese (Chinese speech masking Chinese target sentences). It will be argued that the differences in same-language masking between English and Chinese listeners reflect structural differences between English and Mandarin Chinese. [Work supported by China NSF and CIHR.]

**4pSC11. Cross-linguistic differences in speech perception.** Keith Johnson and Molly Babel (UC Berkeley, 1203 Dwinelle Hall, Berkeley, CA 94720-2650)

This research explores language-specific perception of speech sounds. This paper discusses two experiments: experiment 1 is a speeded forced-choice AX discrimination task and experiment 2 is a similarity rating task. Experiment 1 was intended to investigate the basic auditory perception of the listeners. It was predicted that listeners’ native languages would not influence responses in experiment 1. Experiment 2 asked subjects to rate the similarity between two tokens on a five-point equal interval scale; the purpose of this experiment was to explore listeners’ subjective impression of speech sounds. In experiment 2 it was predicted that listeners’ language would affect their responses. The same stimuli were used in both experiments. The stimuli consisted of vowel-fricative-vowel sequences produced by a trained phonetician. Six fricatives were used: /f, th, s, sh, x, h/. These fricatives were embedded in three vowel environments: /a\_a/, /i\_i/, and /u\_u/. Tokens were presented to listeners over headphones with a 100-ms interval. Independent groups of 15 native Dutch and English listeners participated in each of the two experiments. Results suggest that listeners’ language influenced responses in both experiments, albeit the result was larger in experiment 2. [Work supported by NIH.]

**4pSC12. Neural coding of perceptual interference at the preattentive level.** Yang Zhang (Dept. of Speech-Lang.-Hearing Sci., Univ. of Minnesota, Minneapolis, MN 55455), Patricia Kuhl, Toshiaki Imada (Univ. of Washington, Seattle, WA 98195), Toshiaki Imada, and Masaki Kawakatsu (Tokyo Denki Univ., Inzai-shi, Chiba 270-1382, Japan)

Language acquisition involves neural commitment to language-specific auditory patterns, which may interfere with second language learning. This magnetoencephalography study tested whether perceptual interference could occur at the preattentive level. Auditory mismatch field (MMF) responses were recorded from ten American and ten Japanese adult subjects in the passive oddball paradigm. The subjects read self-chosen books and ignored the sounds. Three pairs of synthetic /ra-la/ syllables were used: one cross-category pair varied only in the third formant (F3), and the other two within-category pairs varied only in the second formant (F2). ANOVA results showed a main effect of acoustic dimension with significant interaction with subject groups ( $p<0.01$ ). As reported earlier, American listeners showed larger but later MMF responses for the F3 change. By contrast, Japanese listeners showed larger and earlier MMFs than Americans for changes in F2. Moreover, Japanese listeners had larger and earlier MMF responses for the changes in F2 as against changes in F3, which was more prominent in the right hemisphere than in the left. These results provided further support for the hypothesis that language experience produces neural networks dedicated to the statistical properties of incoming speech experienced in infancy, which later interfere with second language acquisition.

**4pSC13. Russian and Spanish listener's perception of the English tense/lax vowel contrast: Contributions of native language allophony and individual experience.** Maria V. Kondaurova (Program in Linguist., Purdue Univ., West Lafayette, IN 47907) and Alexander L. Francis (Purdue Univ., West Lafayette, IN 47906)

We examined the influence of listeners native phonology on the perception of American English tense and lax front unrounded vowels ([i] and [ɪ]). These vowels are distinguishable according to both spectral quality and duration. Nineteen Russian, 18 Spanish, and 16 American English listeners identified stimuli from a *beat-bit* continuum varying in nine spectral and nine duration steps. English listeners relied predominantly on spectral quality when identifying these vowels, but also showed some reliance on duration. Russian and Spanish speakers relied entirely on duration. Three additional tests examined listeners allophonic use of vowel duration in their native languages. Duration was found to be equally important for the perception of lexical stress for all three language groups. However, the use of duration as a cue to postvocalic consonant voicing differed due to phonotactic differences across the three languages. Group results suggest that non-native perception of the English tense/lax vowel contrast is governed by language-independent psychoacoustic factors and/or individual experience. Individual results show large variability within all three language groups, supporting the hypothesis that individual differences in perceptual sensitivity as well as the more frequently cited factors of second language education and experience play an important role in cross-language perception.

**4pSC14. An analysis of acoustic deviation manner in spontaneous speech.** Norimichi Hosogai, Kanae Okita, Takuya Aida, and Shigeaki Okawa (Chiba Inst. of Technol., 2-17-1 Tsudanuma, Narashino, Chiba 275-0016, Japan)

Natural speech typically contains various phenomena deviated from the formal mode such as read speech. It is well known that those paralinguistic phenomena have an important role to give the human emotions and the state of the speakers in speech communication. This study attempts to extract the deviation as an acoustic "vagueness," defined by temporal and dynamical acoustic features of speech. Especially the change of the vagueness during a certain period of speech, such as a 10-minute presentation, is focused. As the acoustic features, it used (i) modulation spectrum and (ii) syllable speed, which may have relations to the speech clarity and the tempo. For the experiments, 70 academic presentation speech data in the Corpus of Spontaneous Japanese (CSJ) are used. As the experimental results, significant properties in the patterns of the modulation spectrum and the syllable speed are obtained as a difference of the beginning and the ending periods of the presentation. This result will contribute to a human-like speech dialog system.

**4pSC15. Nondurational cues for durational contrast in Japanese.** Kaori Idemaru (Dept. of East Asian Lang. and Lit., Univ. of Oregon, 1248 Univ. of Oregon, Eugene, OR 97403) and Susan G Guion (Univ. of Oregon, 1290 Eugene, OR 97403)

This study explores potential secondary cues to a durational contrast by examining short and long stop consonants in Japanese. Durational contrasts are susceptible to considerable variability in temporal dimensions caused by changes in speaking rate. In this study, the proposal is examined that multiple acoustic features covary with the stop length distinction and that these features may aid in accessing the percept intended by the speaker, even when the primary cue, closure duration, is unreliable. The results support the proposal, revealing the presence of multiple acoustic features covarying with the short versus long contrast. Not only are there durational correlates to this contrast—the preceding vowel is longer and the following vowel is shorter for geminates than singletons—but there also are nondurational features covarying with this contrast. Greater fundamental frequency and intensity drops are found from the preceding to the following vowel for the geminate than the singleton stops. These results suggest the possibility that systematic variation of these acoustic

features is used in the perceptual categorization of the contrast in addition to the primary cue of closure duration. Moreover, the nondurational correlates are promising candidates for speech-rate resistant features.

**4pSC16. Different motor strategies for increasing speaking rate: Data and modeling.** Majid Zandipour, Joseph Perkell, Mark Tiede, Frank Guenther (M.I.T., Res. Lab Electron., Speech Commun. Group, 50 Vassar St, Cambridge, MA 02139, majidz@speech.mit.edu), Kiyoshi Honda (ATR Human Information Processing Res. Lab., Kyoto 619-0288, Japan), and Emi Murano (Univ. Maryland Dental School, Baltimore, MD, 21209)

Different motor strategies for increasing speaking rate: data and modeling EMG, kinematic and acoustic signals were recorded from two male subjects as they pronounced multiple repetitions of simple nonsense utterances. The resulting data indicate that the two subjects employed different motor strategies to increase speaking rate. When speaking faster, S1 significantly increased the size of the articulatory target region for his tongue movements, increased the speed of the tongue movements and the rate of EMG rise somewhat, while decreasing the movement duration significantly and movement distance slightly. In contrast, at the fast rate, S2 had the same size articulatory target region and rate of EMG rise as at the normal rate, but decreased the speed, distance, and duration of tongue movement slightly. Each subject had similar dispersions of acoustic targets in  $F1-F2$  space at fast versus normal rates, but both shifted target centroids toward the center of the vowel space at the fast rate. Simulations with a biomechanical model of the vocal tract show how modulations of motor commands may account for such effects of speaking rate on EMG, kinematics, and acoustic outputs. [Work supported by NIDCD, NIH.]

**4pSC17. Effect of speaking rate on individual talker differences in voice-onset-time.** Rachel M. Theodore, Joanne L. Miller, and David DeSteno (Dept. of Psych., 125 NI, Northeastern Univ., 360 Huntington Ave., Boston, MA, 02115-5000, r.theodore@neu.edu)

Recent findings indicate that individual talkers systematically differ in phonetically relevant properties of speech. One such property is voice-onset-time (VOT) in word-initial voiceless stop consonants: at a given rate of speech, some talkers have longer VOTs than others. It is also known that for any given talker, VOT increases as speaking rate slows. We examined whether the pattern of individual differences in VOT holds across variation in rate. For example, if a given talker has relatively short VOTs at one rate, does that talker also have relatively short VOTs at a different rate? Numerous tokens of /t/ were elicited from ten talkers across a range of rates using a magnitude-production procedure. VOT and syllable duration (a metric of speaking rate) were measured for each token. As expected, VOT increased as syllable duration increased (i.e., rate slowed) for each talker. However, the slopes as well as the intercepts of the functions relating VOT to syllable duration differed significantly across talkers. As a consequence, a talker with relatively short VOTs at one rate could have relatively long VOTs at another rate. Thus the pattern of individual talker differences in VOT is rate dependent. [Work supported by NIH/NIDCD.]

**4pSC18. Variation in vowel production.** Joseph Perkell, Majid Zandipour, Satrajit Ghosh, Lucie Menard (Speech Commun. Group, Res. Lab. of Electron., Rm. 36-511, M.I.T., Cambridge, MA 02139), Harlan Lane, Mark Tiede, and Frank Guenther (M.I.T., Cambridge, MA 02139)

Acoustic and articulatory recordings were made of vowel productions by young adult speakers of American English—ten females and ten males—to investigate effects of speaker and speaking condition on measures of contrast and dispersion. The vowels in the words *teat*, *tit*, *tet*, *tat*, *tot*, and *toot* were embedded in two-syllable "compound words" consisting of two CVC syllables, in which each of the two syllables comprised a real word, the consonants were /p/, /t/ or /k/, the two adjoining consonants were always the same, the first syllable was unstressed and the second stressed. Variations of phonetic context and stress were used to induce

dispersion around each vowel centroid. The compound words were embedded in a carrier phrase and were spoken in normal, clear, and fast conditions. Initial analyses of *F1* and *F2* on 15 speakers have shown significant effects of speaker, speaking condition (and also vowel, stress, and context) on vowel contrast, and dispersion around means. Generally, dispersions increased and contrasts diminished going from clear to normal to fast conditions. Results of additional analyses will be reported. [Work supported by NIDCD, NIH.]

**4pSC19. Region, gender, and vowel quality: A word to the wise hearing scientist.** Richard Wright (Dept. of Linguist., Univ. of Washington, Box 354340, Seattle, WA 98195-4340, rawright@u.washington.edu), Stephanie Bor, and Pamela Souza (Univ. of Washington, Seattle, WA 98105)

Sociophonetic research has established effects of regional accent and gender on spoken vowels. Many gender differences are due to sociolinguistic factors and thus vary by region. The implications for researchers and clinicians are important: gender variation must be controlled for according to the region of the listener and talker population. Moreover, speech perception stimuli used in research and in clinical applications have limited regional application. This poster illustrates these factors using the Pacific Northwest regional accent. The data, collected for a previous study on hearing aid processing, consist of three repetitions of eight vowels produced in real-word /h\_d/ (or /\_d/) contexts by six males and six females ranging in age from 19 to 60. Formants were measured using an LPC with an accompanying FFT and spectrogram for verification. The results revealed vowel-specific differences in the male and female speech over and above those typically associated with physiologic predictions, and different again from those observed in past studies from different regions. Taken as a whole, these data suggest that speech and hearing researchers should take care in selecting stimuli for general-use speech perception tests. [Work supported by NIDCD training grant (#DC00033) and NIH RO1 (1 RO1 DC006014).]

**4pSC20. Acoustic characteristics of vowels in three regional dialects of American English.** Ewa Jacewicz, Robert Allen Fox, Yolanda Holt (Speech Acoust. and Percept. Labs., Dept. of Speech and Hearing Sci., The Ohio State Univ., Columbus, OH 43210), and Joseph Salmons (Univ. of Wisconsin—Madison, Madison, WI)

Most of the comparative sociophonetic studies of regional dialect variation have focused on individual vowel differences across dialects as well as speaker variables. The present work seeks to define basic acoustic characteristics of entire vowel systems for three different regional variants of American English spoken in southeastern Wisconsin (affected by the Northern Cities Shift), western North Carolina (affected by the Southern Vowel Shift), and central Ohio (not considered to be affected currently by any vowel shift). Three groups of speakers (men and women) aged 20–29 years were recorded from each geographic area defined by two to three counties (creating a highly homogeneous set of speakers). Acoustic measures for the set of 14 monophthongs and diphthongs in /h\_d/ context included vowel space area for each speaker, global spectral rate of change for diphthongized vowels (defined over the first three formant slopes), the amount of frequency change for *F1* and *F2* at two temporal points located close to vowel onset and offset (vector length), and vowel duration. These measures will establish both systemic and vowel inherent characteristics across the three dialects, serving as a basis for future examination of conditioning factors on vowels in chain shifts. Dialectal differences will be discussed. [Work supported by NIH NIDCD R01 DC006871.]

**4pSC21. The rhythmic characterization of two varieties of Portuguese.** Verna Stockmal, Emilia Alonso Marks, Audra Woods, and Z. S. Bond (Ohio Univ., Athens, OH 45701)

As spoken in Europe, Portuguese is said to be stress-timed, while Brazilian Portuguese appears to display characteristics of both stress and syllable timing [P. A. Barbosa, D.E.L.T.A. **16**, 369–402 (2000)]. We employed the Ramus *et al.* metric, based on acoustic-phonetic measurements [Ramus *et al.*, *Cognition* **73**, 265–292 (1999)], to investigate the possibility of distinguishing between the two varieties of the language. Five native speakers of European Portuguese and five native speakers of Brazilian Portuguese recorded the same short prose passage taken from a magazine. The talkers were asked to read at a normal, comfortable rate. The reading time of the passage averaged 60 s, with considerable differences among the talkers. From the vocalic and consonantal intervals, the Ramus metrics, percent vocalic interval and standard deviation of consonantal and vocalic interval, were calculated. The five talkers of the two language varieties differed on the values of these parameters. The values of %V and SD-V showed overlap between the two talker groups, while the BP talkers tended to show lower values for SD-C. Apparently, the rhythmic characterization of the two varieties of the language is not clearly categorical, but rather ranges along a continuum.

**4pSC22. Indexical cues to talker sex and individual talker identity extracted from vowels produced in sentence-length utterances.** Michael J. Berkowitz (Dept. of Psych., 301 Wilson Hall, Vanderbilt Univ., 111 21st Ave. South, Nashville, TN 37203, michael.j.berkowitz@vanderbilt.edu), Jo-Anne Bachorowski (Vanderbilt Univ., Nashville, TN 37203), and Michael J. Owren (Georgia State Univ., Atlanta, GA 30302)

The purpose of this study was to replicate and extend a previous study of indexical cuing [J.-A. Bachorowski and M. J. Owren, *J. Acoust. Soc. Am.* **106**, 1054–1063 (1999)] by including more vowel sounds spoken in more diverse contexts. Specific combinations of acoustic parameters that should represent talker sex and individual talker identity were identified using predictions based on known sources of variation in vocal production-related anatomy. This study utilized 100 recordings of sentence-length utterances, produced by each of 43 male and 44 female undergraduates, as well as 22 stock-phrase recordings produced by these same participants. One of five vowel sounds (/æ, e, i, ə, u/) was isolated from each sentence and analyzed for *F0*, *F1*, *F2*, *F3*, *F4*, vowel duration, jitter, shimmer, and harmonicity. Classification by talker sex was nearly perfect using a combination of cues related to both vocal-fold and vocal-tract anatomy. The accuracy of classification by individual identity depended strongly on cues relating to vocal tract-variation within sex.

**4pSC23. Utterance-final position and projection of femininity in Japanese.** Mie Hiramoto and Victoria Anderson (Dept. of Linguist., Univ. of Hawaii, 1890 East-West Rd., Honolulu, HI 96822)

Japanese female speakers frequent use of suprasegmental features such as higher pitch, longer duration, wider pitch range, and more instances of rising intonation *vis-a-vis* male speakers, is recognized as Japanese women's language (JWL) prosody. However, these features normally co-occur with gender-specific sentence-final particles (SFPs) like the strongly feminine “kashira.” In this study, we examined the use of pitch and duration in utterances without SFPs, to determine whether JWL prosody is a function of SFPs or of utterance-final position. Eight male and eight female native Japanese speakers were instructed to read prepared sentences as though auditioning for a masculine theater role and then as though auditioning for a feminine role. Results indicate that utterance-final position is the projection point of JWL prosody even in the absence of SFPs. The data used for this study show high pitch, wide pitch range, long duration, and rising intonation at utterance-final positions when produced (by both men and women) in the feminine gender role. Conversely, in the masculine gender

role, both men and women avoid the use of such prosodic features, and may even avoid using rising intonation in interrogative sentences, where the tonal grammar calls for it.

**4pSC24. Attitudinal correlate of final rise-fall intonation in Japanese.** Toshiyuki Sadanobu (Kobe Univ., Tsurukabuto 1-2-1, Nada, Kobe, 657-8501, Japan)

Abrupt rise and subsequent fall intonation is common at the end of intonation units in Japanese, but its attitudinal correlate has not been fully elucidated yet. This intonation appears in the literature of the 1960's as politicians' way of speech, and nowadays not only politicians but many speakers including older generations often use it. However, this intonation is stigmatized as childish, and many people devalue it as an unintelligent way of speaking by young people. Where does this great gap between reality and image of this intonation come from? This presentation addresses this problem by focusing on natural conversation of Japanese daily life. The conclusions are as follows: (i) Rise-fall intonation often appears when the speaker talks about high-level knowledge, whereas it disappears when the speaker talks about their personal experience. (ii) Rise-fall intonation at the end of an intonation conveys the speaker's being so occupied with speaking that intonation unit. The childish image comes from the speaker's unawareness of their overall speech because of being occupied with local process. [Work supported by the Ministry of Education, Science, Sport, and Culture, Grant-in-Aid for Scientific Research (A), 16202006, and by the Ministry of Internal Affairs and Communications, SCOPE 041307003.]

**4pSC25. Vowel devoicing in Japanese infant- and adult-directed speech.** Laurel Fais, Janet Werker (Dept. of Psych., Univ. of BC, 2136 West Mall, Vancouver, BC V6T 1Z4 Canada, jwlab@psych.ubc.ca), Sachio Kajikawa, and Shigeaki Amano (NTT Commun. Sci. Labs., Seika-cho, Soraku-gun, Kyoto 619-0237 Japan)

It is well known that parents make systematic changes in the way they speak to infants; they use higher pitch overall, more pronounced pitch contours, more extreme point vowels, and simplified morphology and syntax (Andruski and Kuhl, 1996; Fernald *et al.*, 1989). Yet, they also preserve information crucial to the infants ability to acquire the phonology of the native language (e.g., phonemic length information, Werker *et al.*, 2006). The question examined in this paper is whether information other than phonemic segmental information is also preserved, namely, information concerning the phonological process of vowel devoicing. Devoicing of high vowels between voiceless consonants and word-finally after a voiceless consonant is a regular and well-attested phonological process in Japanese (Shibatani, 1990). A corpus of speech by Japanese mothers addressed to their infants and addressed to another adult was examined, and the degree and frequency with which they apply vowel devoicing in each type of speech was analyzed. Rates of vowel devoicing in speech to adults and infants are compared, accommodations made to infants and to hearing-impaired children are discussed (Imaizumi *et al.*, 1995), and the implications of these accommodations for the acquisition of vowel devoicing by Japanese infants are explored.

**4pSC26. Language and gender differences in speech overlaps in conversation.** Jiahong Yuan, Mark Liberman, and Christopher Cieri (Univ. of Pennsylvania, Philadelphia, PA 19104)

Language and gender differences in speech overlaps in conversation were investigated, using the LDC CallHome telephone speech corpora of six languages: Arabic, English, German, Japanese, Mandarin, and Spanish. To automatically obtain the speech overlaps between two sides in a conversation, each side was segmented into pause and speaking segments, and the overlap segments during which both sides were speaking were time stamped. Two types of speech overlaps are distinguished: (1) One side takes over the turn before the other side finishes (turn-taking type). (2)

One side speaks in the middle of the other side's turn (backchannel type). It was found that Japanese conversations have more short (less than 500 ms) turn-taking type of overlap segments than the other languages. The average number of such segments per 10 min of conversation for Japanese was 62.6 whereas the average numbers for the other languages ranged from 37.9 to 43.3. There were, however, no significant differences among the languages on the backchannel type of overlaps. Cross-linguistically, conversations between two females contained more speech overlaps (both types) than those between a male and a female or between two males, and there was no significant difference between the latter two.

**4pSC27. An acoustic study of laryngeal contrast in three American Indian Languages.** Heriberto Avelino (Dept. of Linguist., UC Berkeley, Berkeley, CA 94720-2650)

A contrast between modal and nonmodal phonation is commonly found in American Indian languages. The use of laryngealized voice has been reported in a number of languages from different linguistic families. This paper investigates the acoustics of laryngealized phonation in three indigenous languages spoken in Mexico, Yalalag Zapotec, Yucatec Maya, and Mixe. These languages differ in terms of the use of other features controlled by action of the larynx, i.e., tone. In Zapotec there is a contrast between high, low, and falling tones; Maya has phonemic high and low tones, whereas Mixe does not present phonemic pitch. The results show that the production of phonemic laryngeal vowels differs from language to language. The data suggest that the specific implementation of laryngealization depends in part on the relationship with contrastive tone. The patterns of the languages investigated provide new evidence of the possible synchronization of phonation throughout the vowel. With this evidence, a typology of modal/nonmodal phonation in phonation-synchronizing languages is proposed.

**4pSC28. The comparison between Thai and Japanese temporal control characteristics using segmental duration models.** Chatchawarn Hansakunbuntheung and Yoshinori Sagisaka (GITI, Waseda Univ., 29-7 Bldg. 1-3-10, Nishi-Waseda, Shinjuku-ku, Tokyo 169-0051, Japan, chatchawarnh@fuji.waseda.jp)

This paper compares the temporal control characteristics between Thai and Japanese read speech data using segmental duration models. The same and the different control characteristics have been observed from phone level to sentence level. The language-dependent and language-independent control factors have also been observed. In phone and neighboring phone level, different characteristics are found. Japanese vowel durations are mainly compensated by only adjacent preceding and following phones, which results from mora timing. Unlike Japanese, Thai vowel durations are affected by two succeeding phones. It can be guessed that the differences come from syllabic structures. In word level, most content words tend to have longer phone durations while function words have shorter ones. In phrase level, both languages express duration lengthening of syllable/mora at the phrase initial and final. For language-specific factors, Thai tones express small alteration on phone duration. The comparisons explore the duration characteristics of the languages and give more understanding to be used in speech synthesis and second-language learning research. [Work supported in part by Waseda Univ. RISE research project of "Analysis and modeling of human mechanism in speech and language processing" and Grant-in-Aid for Scientific Research A-2, No. 16200016 of JSPS.]

**4pSC29. Articulatory settings of French and English monolinguals and bilinguals.** Ian L. Wilson (Univ. of Aizu, Tsuruga, Ikki-machi, Aizu-Wakamatsu City, Fukushima, 965-8580, Japan, wilson@u-aizu.ac.jp) and Bryan Gick (Univ. of BC, Vancouver, BC V6T1Z1 Canada)

This study investigated articulatory setting (AS), a language's underlying posture of the articulators. Interspeech posture (ISP) of the articulators (their position when motionless during interutterance pauses) was used as a measure of AS in Canadian English and Quebecois French.

Optotrak and ultrasound imaging were used to test whether ISP is language specific in monolingual and bilingual speakers. Results show significant differences in ISP across the monolingual groups, with English exhibiting a higher tongue tip, more protruded upper and lower lips, and narrower horizontal lip aperture. Results also show that upper and lower lip protrusion are greater for the English ISP than for the French ISP, in all bilinguals who were perceived as native speakers of both languages, but in no other bilinguals. Tongue tip height results mirror those of the monolingual groups, for half of the bilinguals perceived as native speakers of both languages. Finally, results show that there is no unique bilingual-mode ISP, but instead one that is equivalent to the monolingual-mode ISP of a bilingual's currently most-used language. This research empirically confirms centuries of noninstrumental evidence for AS, and for bilinguals it suggests that differences between monolingual and bilingual modes do not hold at the phonetic level.

**4pSC30. Temporal and spectral variability of vowels within and across languages with small vowel inventories: Russian, Japanese, and Spanish.** Franzo F. LawII (Speech Acoust. and Percept. Lab., City Univ. of New York—Grad. Ctr., 365 Fifth Ave., New York, NY 10016-4309, flaw@gc.cuny.edu), Yana D. Gilichinskaya, Kikuyo Ito, Miwako Hisagi, Shari Berkowitz, Mieko N. Sperbeck, Marisa Monteleone, and Winifred Strange (City Univ. of New York—Grad. Ctr., New York, NY 10016-4309)

Variability of vowels in three languages with small vowel inventories (Russian, Japanese, and Spanish) was explored. Three male speakers of each language produced vowels in two-syllable nonsense words (VCa) in isolation and three-syllable nonsense words (gaC1VC2a) embedded within carrier sentences in three contexts: bilabial stops in normal rate sentences and alveolar stops in both normal and rapid rate sentences. Dependent variables were syllable duration and formant frequency at syllable midpoint. Results showed very little variation across consonant and rate conditions in formants for /i/ in Russian and Japanese. Japanese short /u, o, a/ showed fronting (F2 increases) in alveolar context, which was more pronounced in rapid sentences. Fronting of Japanese long vowels was less pronounced. Japanese long/short vowel ratios varied with speaking style (isolation versus sentences) and speaking rate. All Russian vowels except /i/ were fronted in alveolar context, but showed little change in either spectrum or duration with speaking rate. Spanish showed a strong effect of consonantal context: front vowels were backed in bilabial context and back vowels were fronted in alveolar context, also more pronounced in rapid sentences. Results will be compared to female productions of the same languages, as well as American English production patterns.

**4pSC31. Does infant-directed speech in Tagalog resemble infant-directed speech in American English?** Emmylou Garza-Prisby, Shiri Katz-Gershon, and Jean Andruski (Aud. & Speech-Lang. Pathol. Dept., Wayne State Univ., 207 Rackham, 60 Farnsworth St., Detroit, MI 48202)

This study compared the speech of a Filipino mother in infant- and adult-directed speech in order to investigate whether the mother used the acoustic features typically found in the infant-directed speech of American English-speaking mothers. Little acoustic documentation is available on the acoustic features of Tagalog, and no acoustic comparison of speech registers has so far been conducted. Impressionistically, Tagalog-speaking mothers' do not appear to use the features typically found in American mothers' speech to their young infants. The mother was recorded talking to her infant and to another adult native speaker of Tagalog. Recordings were made in the mother's home and visits occurred during the first 6 months of the infant's life. Specific acoustic features analyzed include (a) vowel duration, (b) vowel formant frequencies, (c) vowel triangle size, (d) rate of speech, (e) fundamental frequency, and (f) F0 range. Morphological and syntactic features were also analyzed, including (g) mean length of utterance and (h) sentence length. Results support a need for further study of speech registers in Filipino mothers.

**4pSC32. Restricting relativized faithfulness and local conjunction in optimality theory.** Haruka Fukazawa (Keio Univ., 4-1-1 Hiyoshi, Kohoku-ku, Yokohama, Japan)

Within the framework of optimality theory (OT), the two mechanisms, relativized faithfulness and local conjunction, play inevitable roles, especially when a simple constraint ranking fails to account for the data. However, their domain of application are too unrestricted and even overlapping each other. For instance, there are some cases which could be explained both by the ranking with relativized faithfulness and by the one with local conjunction. Syllable-final neutralization in German and geminate devoicing in Japanese loanwords are of interest in this context. The present paper proposes formal restrictions mostly on the local conjunction mechanism: the soundness of constraint combination, the number of constraints involved in a conjunction, and the local domain of conjunction. They not only can simplify the analysis but also give a more principled account for the overlapping cases. In fact, relativized faithfulness approach wins over local conjunction approach both in German neutralization and in Japanese loanwords. It is desirable for the universal grammar to be more restricted. Removing an overlap of theoretical devices is an important step toward the goal.

## Session 4pUWa

## Underwater Acoustics: Session in Honor of Leonid Brekhovskikh II

William A. Kuperman, Cochair

*Scripps Inst. of Oceanography, Univ. of California, San Diego, Marine Physical Lab., La Jolla, CA 92093-0238*

Oleg A. Godin, Cochair

*NOAA, Earth System Research Lab., 325 Broadway, Boulder, CO 80305-3328**Invited Papers*

1:00

**4pUWa1. Underwater noise as source of information on conditions and dynamics of ocean environments.** Alexander V. Furduiev (N. N. Andreyev Acoust. Inst., 4 Shvernika St., Moscow 117036, Russia)

Leonid Brekhovskikh wrote in his book *The Ocean and the Human* (1987): "Ocean noise is as important oceanographic parameter as temperature, current, and wind." Brekhovskikh created and headed the Laboratory of Acoustic Methods of Ocean Research in 1956. One of the scientific directions of the Laboratory was investigation of underwater noise both as interference for sound reception and a source of environmental information. Long-term studies on the unique acoustic research vessels created under the initiative of Brekhovskikh resulted in numerous important findings, including ambient noise spectra and envelopes of acoustic fluctuations, depth dependence of noise directivity, and mechanisms of ambient noise generation. Brekhovskikh was always eager to find practical applications of scientific results. Different methods of ensuring noise immunity of hydroacoustic arrays were developed under his supervision. Passive methods of acoustic navigation based on use of natural noise were suggested. Techniques for underwater acoustic monitoring of the ocean based either on ambient noise analysis or reemission of noise from a point away from the receiving system have been developed. The success of the team of scientists headed by Brekhovskikh was determined by the creative atmosphere around him: there was neither competition nor commercial interests. The common goal was knowledge of the ocean.

1:20

**4pUWa2. Distributed acoustic sensing in shallow water.** Henrik Schmidt (Ctr. for Ocean Eng., MIT, Cambridge, MA 02139)

The significance of Leonid Brekhovskikh to the ocean acoustics is undisputed. He was pioneering not only in terms of fundamental understanding of the ocean acoustic waveguide, but also the development of efficient and numerically stable approaches to propagation of sound in a stratified ocean. As such he has been an inspiration to a whole generation of model developers, leading to today's suite of extremely powerful wave theory models, capable of accurately representing the complexity of the shallow-water ocean waveguide physics. The availability of these computational tools have in turn led to major advances in adaptive, model-based signal-processing techniques. However, such computationally intensive approaches are not necessarily optimal for the next generation of acoustic sensing systems. Thus, ocean observation in general is currently experiencing a paradigm shift away from platform-centric sensing concepts toward distributed sensing systems, made possible by recent advances in underwater robotics. In addition to a fully autonomous capability, the latency and limited bandwidth of underwater communication make on-board processing essential for such systems to be operationally feasible. In addition, the reduced sensing capability of the smaller physical apertures may be compensated by using mobility and artificial intelligence to dynamically adapt the sonar configuration to the environment and the tactical situation, and by exploiting multiplatform collaborative sensing. The development of such integrated sensing and control concepts for detection, classification, and localization requires extensive use of artificial intelligence incorporating a fundamental understanding of the ocean acoustic waveguide. No other sources in literature provide this with the clarity and depth that is the trademark of Academician Brekhovskikh's articles and classical textbooks. [Work supported by ONR.]

*Contributed Papers*

1:40

**4pUWa3. When the shear modulus approaches zero: Fluids don't bend and Scholte leaves the room.** Robert I. Odom (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698), Donna L. G. Sylvester, and Caitlin P. McHugh (Seattle Univ., Seattle, WA 98122-1090)

The  $4 \times 4$  linear system of differential equations describing the propagation of the displacements and tractions in an elastic layered medium becomes singular as the shear modulus of the elastic medium approaches zero. There are a number of approximate ways to handle this singularity in

order to impart numerical stability to the computation of the elastic waves in a layered medium. For example, one can impose an irrotational constraint on the displacements or introduce a massive elastic interface (MEI). Both of these ways of handling the weak shear strength are approximate, but avoid the need for singular perturbation theory [Gilbert, 1998]. Scholte waves are interface waves that propagate along the interface between an elastic solid and a fluid. They have nodes near or on the interface and decay exponentially into the bounding media. Scholte waves do not occur at the boundary between fluids. As the shear speed in the bounding elastic medium approaches zero, the Scholte waves disappear from the spectrum. We investigate this disappearance by applying singular perturbation theory

to the coupled fluid-elastic system. Among other things, we will address the rate in wave-number space at which the Scholte waves disappear from the spectrum.

1:55

**4pUWa4. Measurement of the plane-wave reflection coefficient of the ocean bottom and the legacy of Leonid Brekhovskikh.** George V. Frisk (Florida Atlantic Univ., 101 N. Beach Rd., Dania Beach, FL 33004) and Luiz L. Souza (Bose Corp., Framingham, MA 01701)

Leonid Brekhovskikh's classic text [*Waves in Layered Media* (Academic, New York, 1980)] inspired the development of several techniques for measuring the plane-wave reflection coefficient of the ocean bottom. Specifically, his application of the geometrical acoustics approximation to the problem of reflection of a spherical wave from a horizontally stratified medium provided the theoretical foundation for evaluating the strengths and weaknesses of various measurement methods. The most popular method assumes that the reflected field also consists of a spherical wave multiplied by the reflection coefficient evaluated at the specular angle. However, Brekhovskikh's work showed that this interpretation is confined to a limited range of angles and bottom structures and, if applied improperly, can lead to unphysical results such as negative bottom loss. This paper describes a technique which circumvents these deficiencies. It consists of measuring the pressure field magnitude and phase versus range due to a cw point source and Hankel transforming these data to obtain the depth-dependent Green's function versus horizontal wavenumber. The reflection coefficient is then obtained from the Green's function using the analytical relationship between the two quantities. The method is demonstrated using 220-Hz data obtained in a near-bottom geometry in the Icelandic Basin. [Work supported by ONR.]

2:10

**4pUWa5. Field from a point source above a layered half-space; theory and observations on reflection from the seabed.** Charles W. Holland (Appl. Res. Lab., The Penn State Univ., State College, PA)

L. M. Brekhovskikh's book *Waves in Layered Media* has provided decades of graduate students and researchers alike with a comprehensive and enormously useful reference. One topic from that work, reflection from a point source above a plane-layered medium (the seabed), is discussed here. Both theoretical underpinnings and observations of reflection from a homogeneous halfspace, a transition layer with smoothly varying density and velocity profiles, and discrete layered media are considered for various shallow water sediment fabrics. [Work supported by the Office of Naval Research and NATO Undersea Research Centre.]

2:25

**4pUWa6. Plane-wave model and experimental measurements of the directivity of a Fabry-Perot, polymer film, ultrasound sensor.** Benjamin T. Cox and Paul C. Beard (Dept. of Med. Phys. and Bioengineering, Univ. College London, Gower St., London, WC1E 6BT, UK, bencox@mpb.ucl.ac.uk)

Optical detection of ultrasound is popular due to the small element sizes that can be achieved. One method exploits the thickness change of a Fabry-Perot (FP) interferometer caused by the passage of an acoustic wave to modulate a laser beam. This detection method can have greater sensitivity than piezoelectric detectors for sub-millimeter element sizes. The directivity of FP sensors and the smallest achievable effective element size are examined here. A plane-wave model of the frequency-dependent directional response of the sensor, based on Brekhovskikh's work on elastic waves in layered media, is described and validated against experimental directivity measurements made over a frequency range of 15 MHz and from normal incidence to 80 deg. In terms of applications, the model may be used to provide a noise-free response function that can be deconvolved from sound field measurements in order to improve accuracy in high-frequency metrology and imaging applications, or, for example, as a predictive tool to improve sensor design. Here, the smallest achievable effective element radius was investigated by comparing the directivity with that of a rigid circular pressure transducer, and found to be  $\sim 0.9d$ , where  $d$  is the thickness of the FP interferometer. [Funding was provided by the EPSRC, UK]

2:40

**4pUWa7. The interference head wave and its parametric dependence.** Jee Woong Choi and Peter H. Dahl (Appl. Phys. Lab., Univ. of Washington)

The interference head wave that can exist in the presence of a sound-speed gradient in the sediment, is a precursor arrival representing a transition between the first-order head wave and the zeroth-order refracted wave. Using a parabolic equation (PE) simulation, Choi and Dahl [J. Acoust. Soc. Am. **119**, 2660–2668 (2006)] showed how the small shift in the dominant frequency of the interference head wave behaves as a function of the nondimensional parameter  $\zeta$ , which itself is a function of center frequency, gradient, and range. For example, it was shown that the maximum frequency shift occurring in the vicinity of  $\zeta$  equals 2. In this work, we investigate the amplitude and additional spectral properties of the interference head wave and analyze the cause of the frequency shift phenomenon using the ray theory. The limitation on the application of ray method also will be discussed. Finally, the conclusion will be verified by the time-dependent simulation using the RAM PE algorithm. [Work supported by the ONR.]

## Session 4pUWb

## Underwater Acoustics: Session in Honor of Fredrick Fisher

William A. Kuperman, Chair

*Scripps Inst. of Oceanography, Univ. of California, San Diego, Marine Physical Lab., La Jolla, CA 92093-0238*

Chair's Introduction—3:10

*Invited Papers*

3:15

**4pUWb1. FLIP (Floating Instrument Platform): A major Fred Fisher contribution to ocean science.** Fred N. Spiess, Robert Pinkel, William S. Hodgkiss, John A. Hildebrand (Marine Physical Lab., Scripps Inst. of Oceanogr., UCSD 0205, 9500 Gilman Drive; La Jolla, CA 92093-0205, fspiess@ucsd.edu), and Gerald L. D'Spain (UCSD, La Jolla, CA 92093-0205)

Frederick H. Fisher, a loyal and zealous member of the Acoustical Society of America, was an imaginative and effective developer of new techniques for research in both laboratory and seagoing acoustics. Most notable among his contributions was his work in bringing into being and enhancing the usefulness of the spar buoy laboratory, FLIP, from its inception in 1960. Not only did Fred use FLIP in his own research, its existence and many of its ancillary capabilities constituted a base for the seagoing research of others. The authors of this paper have benefited from FLIP's unique capabilities, starting with long-range sound propagation studies in the 1960's and 1970's. FLIP's stability and deep draft structure provided the platform for development of acoustic Doppler techniques for the measurement of ocean currents. Most recently, FLIP has been involved in studies of marine mammal vocalizations and use of multielement arrays to investigate details of shallow-water propagation. Fred's initial studies of sonar bearing accuracy, for which FLIP's construction was funded, and his dedication to advancing FLIP's ability to contribute to ocean science, constitute a legacy that is being utilized today, more than 40 years after FLIP's launching.

3:35

**4pUWb2. Absorption of sound in seawater and ocean ambient noise, the scientific passions of Fred Fisher.** John A. Hildebrand (Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92093-0205)

Fred Fisher made seminal contributions to ocean acoustics in the understanding of the absorption of sound in seawater and ocean ambient noise. Laboratory data and long-range sound propagation data revealed excess acoustic absorption in seawater. Fred Fisher spent much of his scientific career, beginning with his Ph.D dissertation, teasing out the contributions of various seawater components to sound absorption, and his work on this topic set the standard for understanding and modeling these phenomena. Ambient noise is an important aspect of underwater signal detection and is the focus of recent concerns about disturbance of marine organisms. Fred Fisher made important contributions to ambient noise studies by conducting measurements of vertical directionality, thereby testing models for ambient noise production. The value of archival ambient noise data and recent increases in ambient noise will be discussed.

*Contributed Papers*

3:55

**4pUWb3. Fred Fisher's high-pressure work with eyewash and epsom salts.** Christian de Moustier (Ctr. for Coastal & Ocean Mapping, Chase Ocean Eng. Lab, Univ. of New Hampshire, 24 Colovos Rd., Durham, NH 03824)

Starting in 1957 Fred Fisher led research programs devoted to high-pressure measurements related to the physical chemistry of sound absorption in seawater due to magnesium sulfate and other salts. As he put it, he spent his professional lifetime squeezing epsom salt. His interest in the low-frequency anomalous sound absorption in the ocean below 1 kHz led to the discovery of boric acid as the cause of the low-frequency relaxation. This paper is a short review of Fred Fisher's contributions to our knowledge of sound absorption in seawater, based in part on his carefully handwritten lecture notes and numerous low-pressure discussions.

4:10

**4pUWb4. Fred Fisher and research with acoustic vector sensors; Marine Physical Laboratory's vertical array of directional acoustic receivers and ocean noise.** Gerald L. D'Spain and William S. Hodgkiss (Marine Physical Lab, Scripps Inst. of Oceanogr., La Jolla, CA 93940-0701)

Fred Fisher had boundless enthusiasm for all topics acoustic. A chance encounter with him in the hallway usually led to a half-hour discussion of the latest research efforts at the lab and recent results he found exciting. In the 1980s, Fred became interested in the problem of identifying the physical phenomena forming the pedestal about the horizontal in vertical directionality measurements of the deep ocean's low-frequency noise field. Two competing mechanisms had been proposed: downslope conversion of coastal shipping and noise from high latitude winds coupling into the deep

sound channel due to the shoaling of the sound channel axis. The relative contributions of these two mechanisms possibly could be separated if the azimuthal ambiguity of a vertical line array of hydrophones somehow could be broken. Therefore, Fred proposed to build a vertical array of "DIFAR" sensors, which led to the design and construction of the Marine

Physical Lab's Vertical "DIFAR" Array. This talk will reminisce a bit about Fred as well as present some results from an ambient noise experiment conducted in 1992 on the continental shelf using the Vertical DIFAR Array co-deployed with MPL's freely drifting vector sensors, the Swallow floats. [Work supported by ONR and ONT.]

FRIDAY EVENING, 1 DECEMBER 2006

HAWAII BALLROOM, 7:00 TO 10:00 P.M.

### Awards Ceremony

Anthony A. Atchley, President  
*Acoustical Society of America*

Yôiti Suzuki, President  
*Acoustical Society of Japan*

Acknowledgment of Honolulu Local Meeting Organizing Committees

### Presentation of Fellowship Certificates

Anders Askenfeldt  
Sergio Beristain  
Philippe Blanc-Benon  
David A. Conant  
Andes C. Gade  
Anthony W. Gummer  
Charles W. Holland  
Jody E. Kreiman  
Kevin D. LePage

James A. McAteer  
David R. Palmer  
Marehalli G. Prasad  
Hiroshi Riquimaroux  
Peter A. Rona  
Mark V. Trevorrow  
Michael Vörlander  
Joos Vos  
Ben T. Zinn

Science Writing Award in Acoustics for Journalists to Radek Boschetty

Science Writing Award for Professionals in Acoustics to Edwin Thomas

Announcement of 2005 A. B. Wood Medal and Prize to Aaron Thode

Distinguished Service Citation to Thomas D. Rossing  
Silver Medal in Noise to Alan H. Marsh  
Silver Medal in Physical Acoustics to Henry E. Bass  
Silver Medal in Psychological and Physiological Acoustics to William A. Yost  
Wallace Clement Sabine Award to William J. Cavanaugh

Recognition of Acoustical Society of Japan meeting organizers  
Recognition of Acoustical Society of America meeting organizers