Session 5aAB

Animal Bioacoustics: Marine Acoustic Ecology

Jennifer L. Miksis-Olds, Chair
Center for Coastal and Ocean Mapping, Univ. of New Hampshire, 24 Colovos Rd., Durham, NC 03824

Contributed Papers

8:00

5aAB1. A multi-year time series of marine mammal distribution in the Alaskan Arctic. Jessica Crance and Catherine Berchok (AFSC/NMFS/NOAA, National Marine Mammal Lab., 7600 Sand Point Way NE, Seattle, WA 98115, Jessica.Crance@noaa.gov)

The Marine Mammal Laboratory has deployed long-term, moored, passive acoustic recorders at numerous locations in the Alaskan Arctic since the BOEM-funded Chukchi Acoustics, Oceanography, and Zooplankton (CHAOZ) study began in 2010. These instruments, often collocated with oceanographic instruments, collect year-round passive acoustic data (16 kHz sampling rate, ~30% duty cycle). All acoustic recordings (100%) were analyzed using an in-house Matlab-based manual analysis program for 12 different Arctic and sub-Arctic cetacean and pinniped species, as well as for vessel, seismic airgun, and ice noise. Bowhead and beluga distributions showed similar patterns, with bimodal temporal distributions representing the fall and spring migrations. Infrequent gray whale detections (isolated on nearshore recorders) followed a similar, albeit less defined, bimodal distribution. Walrus and bearded seal calls were detected almost year-round, with peak bearded seal calling from April through June. Detections of sub-Arctic species (e.g., humpback, fin, killer, and minke whales) were rare north of 70° N. Correlation of marine mammal distribution with concurrent oceanographic parameters revealed strong associations with oceanographic and prey variables, more specifically, ice concentration and thickness, chlorophyll, wind speed, and currents. This study illustrates the importance of collecting concurrent long-term passive acoustic and oceanographic data in a rapidly changing environment.

8:15

5aAB2. A preliminary acoustic survey of echolocating marine mammals in the Bering Sea. Kerri Seger (CCOM, Univ. of New Hampshire, 9331 Discovery Way, Apt. C, La Jolla, CA 92037, kseger@ucsd.edu), Jennifer Miksis-Olds (CCOM, Univ. of New Hampshire, Durham, NH), and Bruce Martin (JASCO Res., Ltd., Victoria, BC, Canada)

As the Arctic seas rapidly change with increased ocean temperatures and decreased sea ice extent, traditional Arctic marine mammal distributions will be altered, and non-traditionally Arctic species may shift poleward. Arctic species typically include sperm, bowhead, humpback, right, gray, fin, and blue whales; odontocetes, specifically killer and beluga whales; and several pinnipeds species. Their acoustic presence has been documented because they produce relatively low-frequency sounds that are detectable by many common, remotely-deployed recording platforms. Until recently, however, recording constraints of power and storage limited higher sampling rates and prevented the detection of many high-frequency-producing species in the Arctic seas. Such species likely include Baird's, Cuvier's, and Stejneger's beaked whales, as well as Northern right whale and Pacific white-sided dolphins. Using one of the first long-term data sets to record relatively high frequencies in the Bering Sea, signal types similar to those of Baird's and Cuvier's beaked whales, and Pacific white-sided and Northern right whale dolphins have been detected. Their relative acoustic presence from the last several years will be presented, as well as the acoustic presence of other species, like Risso's dolphin, that may be shifting northward into the Arctic seas. [Work supported by ONR.]

8:30

5aAB3. Did humpback whales go missing off Maui, Hawaii? A comparison of song activity between the 2014/15 and 2015/16 breeding seasons. Anke Kügler (Marine Biology Graduate Program, Univ. of Hawaii at Manoa, 2525 Correa Rd. HIG 132, Honolulu, HI 96822, akuegler@hawaii.edu), Marc O. Lammers (Hawaii Inst. of Marine Biology, Univ. of Hawaii at Manoa, Kaneohe, HI), Eden J. Zang (Oceanwide Sci. Inst., Honolulu, HI), Maxwell B. Kaplan, and T A. Mooney (Woods Hole Oceanographic Inst., Woods Hole, MA)

Each winter, thousands of humpback whales (*Megaptera novaeangliae*) migrate from their high latitude feeding grounds in Alaska to mate and calve in the shallow tropical waters around the Main Hawaiian Islands. Population estimates suggest that up to 10,000 animals winter in Hawaiian waters, making up more than half of the total North Pacific stock. However, in the 2015/16 season, anecdotal reports from commercial operators and researchers tell of an unusually low number of whales compared to previous years off the island of Maui. To examine this issue, data from long-term passive acoustic monitoring with autonomous Ecological Acoustic Recorders (EARs) during the 2014/15 and 2015/16 seasons off the west coast of Maui were analyzed using male chorusing levels as a proxy for relative whale abundance. Root-mean-square sound pressure levels (SPLs) were calculated to compare low frequency acoustic energy (0-1.56 kHz) between both seasons. The data showed that chorusing levels dropped in early January 2016. Although levels subsequently increased again until they peaked in February as expected, SPLs remained lower than in 2014/15 throughout the remaining season. These results suggest that the number of singing males was indeed lower during the 2015/16 season off west Maui and that future monitoring is warranted.

8:45

5aAB4. How reliable is song as a cue for acoustic surveys of humpback whales (*Megaptera novaeangliae*) with changing population size and density? Michael J. Noad, Rebecca A. Dunlop, and Amelia Mack (School of Veterinary Sci., The Univ. of Queensland, Gatton, QLD 4343, Australia, mnoad@uq.edu.au)

Acoustic surveys of marine mammals have several advantages over traditional visual surveys. For an acoustic survey to be accurate, however, there needs to be a predictable relationship between vocal behavior and density. Humpback whale's song, produced by males primarily during the breeding season, is potentially a good candidate for such surveys. The function of song, however, remains enigmatic and so its usefulness as a predictor of population density is still not known, particularly when population density changes. We examined aspects of singing behavior during migration off eastern Australia between 1997 and 2015 including the proportion of whales that sang, the spacing between singers, and the use of song by males when joining other whales. Over this period, there was a six-fold increase in the population but this had little effect on the proportion of whales singing, and singers spaced randomly, implying that it could be useful as a survey tool. However, the proportion of singers that continued to sing while joining other whales decreased, showing some density effects on singing behavior. Altogether, while singing behavior can be used as a reasonable proxy of population size at the densities observed, behavioral strategies may start to confound this relationship at higher densities.

5a FRI. AM

9:00

5aAB5. The correlation between the local environment and gray whale behavior as tracked using a sparse hydrophone array in Monterey Bay National Marine Sanctuary. Regina A. Guazzo, John A. Hildebrand, Sean M. Wiggins, Gerald L. D'Spain (Scripps Inst. of Oceanogr., Univ. of California, San Diego, UCSD/SIO/0205, 9500 Gilman Dr., La Jolla, CA 92093-0205, rguazzo@ucsd.edu), and David W. Weller (Southwest Fisheries Sci. Ctr., NOAA, La Jolla, CA)

Gray whales (*Eschrichtius robustus*) make one of the longest annual migrations of any marine mammal. While the seasonal periodicity of the migration and related travel corridors have been well-studied, little is known about how these whales use acoustics while migrating. Acoustic array processing is a powerful tool to localize and track vocalizing whales without disrupting their behavior. Concurrent with the NOAA Southwest Fisheries Science Center visual and infrared gray whale census, four hydrophone-recorder packages with sites separated by approximately 2 km in 58 to 110 m water in Monterey Bay National Marine Sanctuary offshore of Granite Canyon, California, recorded from November 2014 to June 2015. An automatic call detector identified gray whale M3 calls in the dataset. Spectrogram cross-correlation was used to measure the time-difference-of-arrival of these calls and estimate the location of the calling animal. These localizations provide gray whale swimming and calling behavior and indicate changes on diel and seasonal time scales. Results from 15 high-quality tracks show calling gray whales swimming at a mean speed of 1.97 m/s with a range of 1.18 to 2.54 m/s. Acoustic detections and tracks are compared with environmental parameters to investigate how gray whale behavior relates to their physical environment. These results are necessary for understanding the cues gray whales use to navigate from their feeding grounds offshore of Alaska to their breeding grounds in Baja California.

9:15

5aAB6. Vocalization source level distributions and pulse compression gains of diverse baleen whale species in the Gulf of Maine. Delin Wang, Wei Huang, Heriberto Garcia (Elec. and Comput. Eng., Northeastern Univ., 006 Hayden Hall, 370 Huntington Ave., Boston, MA 02115, wang.del@husky.neu.edu), Nicholas C. Makris (Mech. Eng., Massachusetts Inst. of Technol., Cambridge, MA), and Purnima Ratilal (Elec. and Comput. Eng., Northeastern Univ., Boston, MA)

The vocalization source level distributions and pulse compression gains are estimated for four distinct baleen whale species in the Gulf of Maine: blue, fin, minke, and sei. The vocalizations were received on a large-aperture densely sampled coherent hydrophone array system deployed to monitor marine mammals over instantaneous wide areas using the passive ocean acoustic waveguide remote sensing technique. For each baleen whale species, between 400 to over 1400 measured vocalizations with significantly high signal-to-noise ratios (SNR$>$10 dB) after coherent beamforming and localized with high accuracies ($<$10% localization errors) over ranges spanning roughly 1 km to 38 km are included in the analysis. The whale vocalization received pressure levels are corrected for broadband transmission losses modelled using a calibrated parabolic equation based acoustic propagation model for a random range-dependent ocean waveguide. The broadband vocalization equivalent pulse-compression gains are found to be 2.5 ± 1.1 for fin, 24 ± 10 for blue, and 69 ± 23 for sei. The source level distributions and pulse compression gains are essential for determining signal-to-noise ratios and hence detection regions for baleen whale vocalizations received passively on underwater acoustic sensing systems, as well as for assessing communication ranges and potential for echolocation in baleen whales.

9:30

5aAB7. Blind channel estimation of time-varying underwater acoustic waveguide impulse responses. Brendan P. Rideout, Eva-Marie Nosal (Dept. of Ocean and Resources Eng., Univ. of Hawaii at Manoa, 2540 Dole St., Holmes Hall 402, Honolulu, HI 96822, bprideou@hawaii.edu), and Anders Host-Madsen (Elec. Eng., Univ. of Hawaii at Manoa, Honolulu, HI)

Some techniques for underwater passive acoustic localization make use of estimates for the direct and/or interface-reflected acoustic arrival times of a source at one or more underwater hydrophones. This estimation task can be difficult for non-impulsive sources (e.g., humpback whales) due to early arrivals masking later ones. In linear system analysis, the impulse response (IR) of a system is the system output when the input is an impulse (i.e., a short duration, large bandwidth signal), and expresses how a source signal interacts with the environment to yield the output waveform (e.g., that recorded by a hydrophone). Recordings of a relatively impulsive vocalization (e.g., a walrus knock), given that they approximate the IR between the walrus and hydrophone, may facilitate arrival time estimation. IR estimation for unknown, non-impulsive calls is more challenging, particularly for time-varying channels. Blind channel estimation is the process of estimating the set of IRs between a single (unknown) source and multiple receivers, and can potentially help estimate direct and interface-reflected arrival times for non-impulsive marine mammal vocalizations since the IRs can be analyzed rather than waveforms/spectra. In this paper, we use simulations to explore the importance of ocean surface gravity waves on blind estimation of acoustic IRs.

9:45–10:00 Break

10:00

5aAB8. Low-frequency propagation and noise in the intertidal zone. James S. Martin, Peter H. Rogers (School of Mech. Eng., Georgia Inst. of Technol., 771 Ferst Dr., Love Bldg., Rm. 115, Atlanta, GA 30332-0405, james.martin@me.gatech.edu), Ashwin Bhandiwad, Orphal Colleye, and Joseph Sisneros (Dept. of Psych., Univ. of Washington, Seattle, WA)

Measurements of ambient noise and short-range propagation ($<$2m) were made in the intertidal zone with water depths ranging from 50 to 80 cm using vector sensors. The sites selected for the measurements were known nesting areas for plainfin midshipman (Porichthys notatus). Males of this species produce continuous acoustic advertisement calls with fundamental frequencies in the 70 to 100 Hz range. The temporal and spectral structure of these calls must convey information regarding the source's range, direction, and/or fitness over the maximum detectable range in a complex environment. Both fundamental and harmonic components appear to be above a rapid low-frequency increase in the noise floor due to local wave motion and below the frequency range of several observed noise transients at times with modest wind speeds ($<$5 mph) and small wave heights ($<$20cm). There is little evidence of significant propagation differences between specific frequencies within the band of interest.

10:15

5aAB9. Passive acoustic monitoring of fishes and a crustacean using a towed hydrophone in Tateyama bay. Ryuzo Takahashi (Fisheries Technol. Dept., Japan Fisheries Res. and Education Agency, 7620-7, Hasaki, Kamisu-shi, Ibaraki-ken 314-0408, Japan, takahashiryuzo@affrc.go.jp), Tomonari Akamatsu (Japan Fisheries Res. and Education Agency, Yokohama-shi, kanagawa-ken, Japan), Tomohito Imaizumi (Fisheries Technol. Dept., Japan Fisheries Res. and Education Agency, Kamisu-shi, Ibaraki-ken, Japan), and Ikuo Matsuo (Tohoku Gakuin Univ., Sendai-shi, Miyagi-ken, Japan)

Passive acoustic monitoring is useful to observe presence of marine organisms such as marine mammals, fishes and crustaceans. We aimed to estimate the distribution of fishes and crustaceans using a towed hydrophone. A range-wide passive acoustic monitoring was conducted in Tateyama bay in Japan using R/V Takamaru (61t) on July, September and November in 2015 of day and night. The towed hydrophone (Towed Aquafeeler, AquaSound Inc., Japan) was used for the observation. The vessel cruised 5 knot on 5 lines to cover the focal survey area (N 35°020'—N 35°02' and E 139°46—E 139°51). Recorded data with frequency from 100 Hz to 20 kHz were analyzed. Recorded fish sounds have dominant frequency between 200 Hz and 1500 Hz and duration between 0.1 s and 2.1 s. Crustacean sound shows dominant frequency as 3 kHz–20 kHz with the duration of 10 ms—20 ms. Using these features, number of fish sounds was counted automatically using custom-made MATLAB program. Large number of the sounds were recorded in the nighttime than in the daytime. Most frequently received month was July. Both of the fish and crustacean sounds were observed in shallow water ($<$30m) showing acoustic presence map.

10:30

5aAB10. Seasonal acoustic presence of fin and bowhead whales in relation to prey abundance and oceanographic environments in the southern Chukchi Sea. Koki Tsujii, Mayuko Otsuki (Graduate School of Environ. Sci., Hokkaido Univ., 20-5 Benten-cho, Hakodate, Hokkaido 040-0051, Japan, kou1114t@gmail.com), Tomonari Akamatsu (National Res. Inst. of Fisheries Sci., Fisheries Res. Agency, Yokohama, Japan), Ikuo Matsuo (Tohoku Gakuin Univ., Sendai, Japan), Kazuo Amakasu (Tokyo Univ. of Marine Sci. and Technol., Konan, Japan), Minoru Kitamura, Takashi Kikuchi (Japan Agency for Marine-Earth Sci. and Technol., Yokosuka, Japan), Kazushi Miyashita, and Yoko Mitani (Field Sci. Ctr. for Northern Biosphere, Hokkaido Univ., Hakodate, Japan)

To assess how seasonal and ice-associated whales differently respond to the Arctic climate changes, understanding the relationships among their presence and environmental conditions is necessary. We examined the seasonal acoustic presence of fin *Balaenoptera physalus* and bowhead whales *Balaena mysticetus* in relation to prey abundance and oceanographic conditions in the southern Chukchi Sea from July 2012 to July 2014 using passive and active acoustic methods. Oceanographic conditions obtained were water temperature, salinity and sea ice concentrations. Fin whale calls were detected from summer to fall, only in ice-free and high prey abundance period. Moreover, the call-detected period was longer in 2013 than 2012. In contrast, bowhead whale calls were mainly detected from fall to winter and in spring, during ice-free and ice-covered periods. The call presence of bowhead whales overlapped less frequently with that of fin whales in fall 2013 than 2012. The reason for the differences in their acoustic presence between these two years may be caused by higher water temperature and late sea-ice formation in 2013 compared to those in 2012. These results suggest that annual variations in the oceanographic conditions possibly affect the distribution of both species in the southern Chukchi Sea.

10:45

5aAB11. Presence of ribbon seal vocalizations are related to sea ice extent in the Nemuro Strait, the Okhotsk Sea. Mayuko Otsuki (Graduate School of Environ. Sci., Hokkaido Univ., 20-5 Benten, Hakodate, Hokkaido 040-0051, Japan, motsuki@ees.hokudai.ac.jp), Tomonari Akamatsu (National Res. Inst. of Fisheries Sci., Japan Fisheries Res. and Education Agency, Yokohama, Japan), Takahiro Nobetsu (Shiretoko Nature Foundation, Rausu, Japan), and Yoko Mitani (Field Sci. Ctr. for Northern Biosphere, Hokkaido Univ., Hakodate, Japan)

Vocalizations of ice-breeding ribbon seals *Histriophoca fasciata* were recorded using underwater passive acoustic methods from November 2012 to March 2014 off the Nemuro Strait, Japan. Seal presence in the strait was examined in relation to the sea ice extent in the Okhotsk Sea. Ribbon seal downsweeps were only detected when sea ice was present in the strait (February and March), with more detections in March leading up to the spring breeding season. Since ribbon seals require ice for breeding, underwater communication for breeding could be needed during the sea ice presence in this strait. Northeasterly winds were another indirect driver of ribbon seal occurrence, since winds from this direction likely transport sea ice from the central Okhotsk Sea into the Nemuro Strait. Downsweep detections decreased in the middle of the day, which is consistent with observations that seals hauled out on the ice during this time, and thus were producing fewer underwater vocalizations. Our results suggest that a decrease in the sea ice extent in the Okhotsk Sea may change the distribution of ribbon seals and impact their breeding behaviour since the Nemuro Strait region of the Okhotsk Sea is likely the southern limit of their breeding range.

11:00

5aAB12. Comparison between trawl volume and composition and acoustic backscatter. Adrienne M. Copeland (Univ. of Hawaii at Manoa, P.O. Box 1346, Kailua, HI 96734, acopelan@hawaii.edu), Whitlow W. Au (Hawaii Inst. of Marine Biology, Kailua, HI), and Jeffrey Polovina (PIFSC, NOAA, Honolulu, HI)

Active acoustics is an important tool to quantitatively assess the densities of pelagic organisms. While it is an important tool, it is hard to determine the organism composition recorded acoustically especially in oligotrophic waters where mixed organism assemblages are common. One such mixed stock in the Hawaii Islands is the Deep Scattering Layers (DSL). These layers are made up of micronekton (small squid, fish, and crustaceans) which may be an important component linking primary producers with higher trophic levels. Trawling can identify the composition in these layers but this tool can be costly and unable to sample as much area as active acoustics. Many organisms in the DSL diurnally migrate to the surface between 0 and 200 m to feed. We surveyed these organisms using active acoustics and trawling collected in June 2013 and March 2014 on the R/V Oscar Elton Sette. We compared the composition of organisms, volume by functional group, and total volume in the trawl with the acoustic backscatter to determine the relationship between trawl and active acoustic data. We determined that the acoustic densities changed as a function of the trawl composition and volume potentially allowing for the ability to exclusively use acoustics in the future.

11:15

5aAB13. Studying deep diving odontocetes potential prey fields' densities and size structure in Hawaii using a DIDSON imaging sonar. Giacomo Giorli and Whitlow W. Au (Hawaii Inst. of Marine Biology, Univ. of Hawaii at Manoa, PO BOX 1306, Kaneohe, HI 96744, giacomog@hawaii.edu)

Deep diving odontocetesodontocetes feed in the deep ocean on mesopelagic squid and fish. The distribution of these prey are poorly known but does affect the distribution of deep diving odontocete predators. Hence, it is important to collect data on the density and composition of potential prey to identify deep sea habitats that attract deep diving odontocetes. We used a DIDSON imaging sonar to investigate the density and size structure of mesopelagic animals along the Kona coast of Hawaii, where beaked, sperm and pilot whales are regularly seen. Data were collected at multiple stations during three research cruises by lowering the sonar in the water column to sample at different depths and monitoring the mesopelagic biomass with an EK60 38 kHz echosounder. Foraging activity of whales was monitored at few stations using passive acoustic recorders and a towed array of hydrophones. We were able to count and size 7068 pelagic animals. Density and size varied in space, with higher densities and bigger animals at the offshore stations. The DIDSON was effective at detecting targets bigger than 10 cm, a size range that includes potential prey of deep diving odontocetes. Preliminary data suggest differences may exists in habitat selection by each deep diving odontocetes.

11:30

5aAB14. Snapping shrimp: Changes in environmental noise in Laguna de Términos, Campeche, México, after hurricanes. Carmen Bazúa Durán (UNAM, Facultad de Ciencias, Circuito Exterior s/n, Ciudad Universitaria, México, D.F. 04510, Mexico, bazua@unam.mx), Pamela A. Azamar Reyes (FES Iztacala, UNAM, Tlalnepantla, Edo. de México, Mexico), and Whitlow W. Au (MMRP, HIMB, Univ. of Hawaii, Kaneohe, HI)

Laguna de Términos, Campeche, México, is one of the largest coastal lagoons in Mexico where snapping shrimp is the main source of biological sound. We measured the environmental noise from 2004 to 2008 in stations distributed homogeneously in this lagoon during 17 sampling periods using a digital audio tape recorder sampling at 48 kHz with 16 bits for 1 minute at each station. The environmental noise in the recordings was measured with a semi-automatic MATLAB routine designed for this purpose. Results indicate that environmental noise in the lagoon changes from one sampling to another, but it changed drastically after the hurricanes of 2005 in some of the stations. This changes in submarine environmental noise helps us in understanding the effects of natural phenomena in the distribution and abundance of marine wildlife, such as snapping shrimps and bottlenose dolphins. [Work supported by CONACyT-Campeche and PAPIIT&PASPA-UNAM.]

5a FRI. AM

Session 5aAO

Acoustical Oceanography: Topics in Acoustical Oceanography II

Boris Katsnelson, Cochair
Marine Geosciences, University of Haifa, Mt. Carmel, Haifa 31905, Israel

Contributed Papers

8:30

5aAO1. Effect of sand and silt particles on the attenuation of compressional waves in marine mud sediments. Allan D. Pierce (Retired, PO Box 339, 399 Quaker Meeting House Rd., East Sandwich, MA 02537, allanpierce@verizon.net), William L. Siegmann, and Elisabeth Brown (Mathematical Sci., Rensselaer Polytechnic Inst., Troy, NY)

Mud in marine sediments is a mixture of clay, sand, and silt particles. Present paper follows up on a suggestion by Holland and Dosso (JASA, 2013) that the variability of the measured frequency-dependent compressional wave attenuation may be caused by the variability of the amounts of sand and silt particles. The premise is that the porosity for the mud is high and that the sand and silt particles are in suspension. They do not settle out to the bottom of the layer because the card-house fabric of the clay particles tends to hold them in place. This supposition leads to a theory where the clay configuration gives a base-line attenuation, and the contribution from the individual sand and silt particles is additive. The estimation of the latter is distinguished from the existing theories of attenuation of sound in sandy/silty sediments in that the particles are presumed not to touch each other. Particles are assumed to be spherical and there is no slip between particle surfaces and the surrounding water. Earlier theories of attenuation in suspensions by Lamb (Hydrodynamics), Urick (JASA. 1948), and others are criticized because of their assumption that vorticity in fluid is zero. The present theory predicts that the attenuation contribution from a given category of particle is proportional to the square of the frequency and to square of particle diameter, and inversely proportional to the viscosity, in the limit of low frequencies. It approaches a frequency-independent constant at higher frequencies. [Work supported by ONR.]

8:45

5aAO2. Acoustic wave propagation analysis for detecting a sensor robot in heat transfer pipe. Yoshihiro Nishimura, Yoshinobu Kawagoe, and Yoshihisa Harada (AMRI, AIST, 2-1-2, Namiki, Tsukuba, Ibaraki 3053564, Japan, ykawagoe7@gmail.com)

Ultrasonic Testings of heat transfer pipes in electric power plant are important. Those pipes, which are over 60 m and whose inner diameters are 25 mm, 50 mm, and 75 mm, have more than 20 curves and are located at directly unaccessible place. We developed a cableless robot for measuring thickness of pipe in a pipe filled with water but it is difficult to know where our robot is in the pipe at real time. We developed a projector to propagate an intensive single-mode wave into the pipes and examined the relations between frequencies-intenseness and wave propagation (3 kHz–50 kHz). Attenuations of acoustic waves in the pipe filled with water were much lager than those in the pie filled with air and examined for metal pipes and plastic pipes. Acoustic wave tunneling through pipe-wall was calculated and compared with experimental results. Acoustic wave tunneling through the pipe-wall was found to be the main reason why acoustic wave could not propagate longer in the pipe filed with water than that with air. We derived useful data to develop a system for measurement of where our robot is.

9:00

5aAO3. On using multi-look synthetic aperture sonar analysis for the investigation of scattering mechanisms. Anthony P. Lyons (Ctr. for Coastal and Ocean Mapping, Univ. of New Hampshire, University of New Hampshire, Durham, NH 03824, anthony.lyons@ccom.unh.edu) and Daniel C. Brown (Appl. Res. Lab., Penn State Univ., State College, PA)

Multi-look coherence explores the information content of synthetic aperture sonar (SAS) images by splitting the total angle and frequency spectral bandwidth of a complex SAS image into sub-bands. In the presence of a random distribution of surface or volume scatterers, the spectral coherence is proportional to the degree of overlap in sub-bands of the SAS image. Consequently, for a perfectly random distribution of scatterers, the spectral coherence should be zero when the sub-bands do not overlap. If the scatterer distribution deviates from this basic random distribution, the spectral coherence may be non-zero for special cases that would be related to scattering from specific structures such as points or facets. For that reason it appears to be possible to separate coherent features from the random scattered background signal using spectral coherence, suggesting possibilities for detecting and/or characterizing objects (especially in low signal-to-reverberation situations). In this talk, we will discuss the concept and the utility of using the complex coherence between a series of sub-look SAS images to investigate different scattering mechanisms as well as show examples of the technique using real and modeled SAS data. [Work performed under ONR Grants N00014-16-1-2469 and N00014-16-1-2313.]

9:15

5aAO4. Multi target positioning system using Gold code pingers applied to Japanese squid in Japan Sea. Keiichi Uchida (Tokyo Univ. of Marine Sci. and Technol., 5-3-6, ko-nan, Minato-ku, Tokyo 108-8477, Japan, kuchida@kaiyodai.ac.jp), Toyoki Sasakura (AquaSound, Inc., Tokyo, Japan), Takashi Kamoshida (AquaSound, Inc., Kobe, Japan), Aki Miyagi (AquaSound, Inc., Tokyo, Japan), Yasuo Fujikawa (NICHIA Corp., Yokohama, Japan), Keiichi Kamiya (HAMADE Co., Ltd., Hakodate, Japan), Hiroshi Inada, and Yoshinori Miyamoto (Tokyo Univ. of Marine Sci. and Technol., Tokyo, Japan)

To study the effect of light on Japanese squid (*Todarodes pacificus*) behavior, we developed a new ultrasonic biotelemetry system that can simultaneously observe and measure multiple targets. We tested the system on 40 squids in open water. Each squid was attached with a Gold code pinger (AquaSound Inc.) that transmits single frequency signals of multiple codes (IDs). By using the combination of 32 kinds of Gold code sequences, we succeeded to develop a receiving system that can simultaneously track 40 pingers. In the field test, 40 squids attached with the Gold code pingers were released simultaneously into open water and signals were recorded from three fishing boats for a maximum of six hours. The underwater positioning of the squids was calculated based on the "reverse cross-bearing" method by using the cross points from each direction against the target obtained by two boats. With the addition of depth data obtained by the pressure sensor installed in each Gold code pinger, we succeeded to measure the 3D positions of all 40 target squids. Considering the movement of boats by wind and tide, the positions of all three boats were measured every second using the GNSS (FURUNO Electronics Co., Ltd.).

5aAO5. Acoustic characterization of the new Arctic during Ice Exercise 2016. John E. Joseph, D. B. Reeder, and Mitchell S. Nelson (Oceanogr., Naval Postgrad. School, 833 Dyer Rd., Monterey, CA 93943, jejoseph@nps.edu)

The Naval Postgraduate School (NPS) participated in Ice Exercise 2016 (ICEX-16), a multi-national naval exercise conducted in the Beaufort Sea during March 2016. Operating at a remote ice camp, NPS deployed several conductivity, temperature and depth (CTD) sensors to capture oceanographic variability to 500 m depth while performing a series of propagation tests. Mobile and dipped mid-frequency sources transmitted signals to a pair of vertical line array receivers positioned in the field to investigate depth, range, angular and specular characteristics of acoustic propagation and their correlation to variability in oceanographic structure and under-ice conditions. CTD data indicate there was significant variability in sound speed at 50m depth where cold, fresh mixed-layer water interfaces with contrasting warm, saline Pacific Summer Water (PSW) that lays immediately below it. The data also show a persistent and stable subsurface sound channel existed as a result of the PSW with peak temperature at 80 m situated above colder Pacific Winter Water (PWW), resulting in a sound channel axis near 140 m depth. Both features have important implications on sonar performance in the Arctic. Modeled and measured transmission loss are compared to quantify the effects.

5aAO6. Modeling reflection and scattering by deep-ocean turbidites. Darrell Jackson and Dajun Tang (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, drj@apl.washington.edu)

Turbidites are commonly found in deep ocean sediments and are composed of multiple thin layers of fine and coarser sediments, the coarser sediments being deposited by turbidity currents. Although acoustics is a unique means to quantitatively probe the physical and geophysical properties of turbidites, research on this subject is still limited. The work of Gilbert and of Holland and Muncill has shown that this layering can cause acoustic reflection to be very strong, with bottom loss less than that expected for hard seafloors. We examine this problem further, looking at the effect of layering on penetration into the seafloor and also considering scattering by roughness of the inter-layer interfaces. It is found that penetration can be severely limited in some cases, and that this behavior can be understood as a form of Anderson localization. Prospects for inversion of broadband, vertical-incidence data are considered. [Work supported by ONR.]

10:00–10:15 Break

5aAO7. Doppler sonar measurements of bedload sediment transport: Yes there's a signal, but can it be quantified? Len Zedel (Phys. and Physical Oceanogr., Memorial Univ. of NF, Chemistry-Phys. Bldg., St. John's, NF A1B 3X7, Canada, zedel@mun.ca), Jenna Hare, Alex Hay, and Greg Wilson (Oceanogr., Dalhousie Univ., Halifax, NS, Canada)

It has been demonstrated that Doppler sonar measurements of apparent bottom velocity in rivers is correlated with bedload transport. We explore the degree to which this measurement can be quantified using a field deployable multi-frequency (1.2-2.2 MHz), bistatic Doppler sonar system that provides three-component velocity profiles over a ~30 cm interval with ~5 mm resolution at a rate of 50 profiles/sec. Model simulations of system performance demonstrate that the estimates of bed movement are proportional to sediment flux but scaling to actual sediment flux would require an empirical fit. In order to explore the system capability under actual field conditions, a series of experiments was undertaken at the St. Anthony Falls Laboratory (SAFL). The SAFL facility provides a 1.8 m deep, 2.75 m wide flume tank that allows flow rates of order 1 m/s over a mobile bed and where the bedload transport can be measured by sediment traps built into the flume system. We report on preliminary results from trials made at SAFL.

5aAO8. Sound intensity fluctuations as evidence of mode coupling due to moving nonlinear internal waves in shallow water. Boris Katsnelson (Marine GeoSci., Univ. of Haifa, 199 Abba Khouchy Ave., Haifa 3498838, Israel, bkatsnels@univ.haifa.ac.il) and Yun Ren (State Key Lab, Inst. of Acoust., CAS, Beijing, China)

One of manifestation of mode coupling produced by moving nonlinear internal (NIW) waves is temporal fluctuations of intensity or amplitude. In spectrum of fluctuations measured during a few hours there are maximums at frequencies, corresponding to periodical (in time) interaction of moving NIW with periodical (in range) interference structure of the sound field. If velocity of NIW along acoustic track is v, then peaks are placed at frequencies $F_{mn} \sim v(k_m - k_n)$, where k_m is propagation constant of mode m. Due to a few remaining modes in a long range propagation, spectrum should have peaks at a few predominating frequencies. In given work on the base of experimental data of ASIAEX 01 and the corresponding theoretical analysis, it is shown existence of predominating frequencies in spectrum of the sound intensity fluctuations in the presence of NIW moving along acoustic track of the length ~35 km. Both speed of NIW and scales of interference beating determine mentioned frequencies. Note that acoustic track consists of two parts with different depths: ~250 m (NIW are moving during time interval considered) and 120 m (position of receiving array). So, spectrum of fluctuations depends on modal structure both parts of this acoustic track. [Work was supported by ISF and NSFC.]

5aAO9. Mapping the acoustical properties of the upper sedimentary layer in Lake Kinneret. Boris Katsnelson (Marine GeoSci., Univ. of Haifa, Mt. Carmel, Haifa 31905, Israel, bkatsnels@univ.haifa.ac.il), Andrey Lunkov (General Phys. Inst., RAS, Moscow, Russian Federation), Regina Katsman (Marine GeoSci., Univ. of Haifa, Haifa, Israel), and Ilia Ostrovsky (Igal Alon Kinneret Lab, IOLR, Haifa, Israel)

A series of field experiments were conducted to estimate the effective acoustical properties of the bottom in Lake Kinneret (Israel). A wideband sound source (300 Hz to 2 kHz) was positioned at different ranges (1 m to 6 km) from a single hydrophone. Both local and long-range measurements of the sound field were carried out. All the acoustic data demonstrated a high reflectivity of the lake bottom at all grazing angles that is specific feature of gas-bearing sediments. Experimental sound pulses were compared with the simulated ones to obtain the values of effective sound speed in the sedimentary layer at different locations in the lake. Using a basic three component sediment's model (water, rigid particles, and gas bubbles), the gas fraction in the bottom was estimated. A high correlation between the calculated gas fraction and previously measured organic matter concentration was observed. [Work was supported by ISF.]

5aAO10. Model-based design of multibeam sonar system. Jun Kuroda (Radio Application, Guidance and Electro-Optics Div., NEC Corp., 10, Nisshincho 1-chome, Fuchu City, Tokyo 211-8666, Japan, j-kuroda@bq.jp.nec.com) and Motohiro Kaname (Maritime Sensor Eng. Div., NEC Network and Sensor Systems, Ltd., Fuchu City, Tokyo, Japan)

Multibeam sonar is a technique used to many applications in offshore activities, e.g. oceanographic survey, fish finder and sound navigation in maritime. Sonar consists of integrated system of components, which are acoustic transducers, electrical transmitter receivers, and signal processing circuits. Numerical prototyping by model-based design (MBD) is widely employed to develop of various instruments, which are also useful to design algorithm and configuration of sonar. A sonar system is divided roughly into two system categories, that is to say, category 1: acoustic transducers and acoustic fields, and category 2: electrical transmitters, receivers, and signal processing circuits. Design of the category 1 needs numerical simulation of sonic waves reflected by targets. One issue for MBD is to study a high speed and low-cost computation method to simulate echo signals. Numerical computation methods, such as finite-element method (FEM), boundary-element method (BEM) and finite-difference-time-domain method (FDTD), require lots of computation resources to solve large acoustic field. Multibeam sonar is usually used to targets which are away from transducers.

Therefore, equations to simulate transmitted and received sonic wave can be simplified by the far field approximation. This paper proposes a numerical simulation method of acoustic echo signals between transducers and targets for MBD.

11:15

5aAO11. Ultrasonic characterization of seagrass leaf blades (*Posidonia oceanica*). Jay R. Johnson, Gabriel R. Venegas, Preston S. Wilson (Mech. Eng., The Univ. of Texas at Austin, 1 University Station, C2200, Austin, TX 78712, johnson.jayrichard@utexas.edu), Jean-Pierre Hermand (LISA Environ. HydroAcoust. Lab, Université libre de Bruxelles (ULB), Brussels, Brussels Capital, Belgium), and Thibault Urban (NDT Dept., Vinçotte, Vilvoorde, Belgium)

It has been shown that simple mixture theory models are inadequate to describe low-frequency sound propagation through seagrass meadows and further understanding of the acoustic properties of seagrass tissue is necessary. To that end, we present two ultrasonic sound speed measurements on the Mediterranean seagrass *Posidonia oceanica*. First, a transit time measurement through a stack of 120 leaf blades at 2.25 MHz are compared to similar measurements made on macroalgae *Ecklonia radiata*. Sound speed differences are related to the gas content within the aerenchyma of seagrass and other tissue characteristics. Second, ultrasonic sound speed measurements through a suspension of finely divided *P. oceanica* leaf blades at frequencies between 1 and 4 MHz are compared to low-frequency (1-8 kHz) sound speed obtained on the same blade suspension by means of a resonance chamber.

11:30

5aAO12. Simulation of acoustic backscattering from bubbles and droplets under different shape regimes with implications for underwater detection of leakages using active acoustic sensors. Geir Pedersen (Instrumentation, Christian Michelsen Res. AS, P.O. Box 6031, Bergen 5892, Norway, geir.pedersen@cmr.no)

Safe subsea production of oil and gas, as well as storage of CO_2 in geological formations subsea, require rapid detection of accidental releases, eliminating potential harm to the environment. Monitoring of natural gas seeps are important in a climate and environmental aspect. Previous studies demonstrate the potential of sonars for detection and quantification of oil and gas releases of natural or anthropogenic origin. Simulations of acoustic backscattering from plumes is an important tool for design of leakage detection systems and *in situ* leakage quantification. Backscattering from subsea leakages are often simulated assuming spherical shapes using, e.g., effective medium theory, however in real world situations bubble and droplets are not spherical. In this study the backscattering from single bubbles (CO_2, CH_4, and air) and droplets (light and heavy oil) are simulated under different shape regimes (i.e. as a function of Reynolds and Eötvös number). The shapes are extracted from time dependent Computational Fluid Dynamics simulations and the backscattering is modeled using adaptive cross approximation accelerated Boundary Element Method (BEM). The backscattering from single bubbles and droplets are used as input to effective medium theory simulations of plumes. Simulations using BEM and effective medium theory with spherical and non-spherical bubbles and droplets are compared with controlled *in situ* measurements of backscattering with concurrent shape measurements.

11:45

5aAO13. Determination of acoustic waveguide invariant using ships as sources of opportunity in a shallow water marine environment. Christopher M. Verlinden, Jit Sarkar (Marine Physical Lab., Scripps Inst. of Oceanogr., Univ. of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0701, cmverlin@ucsd.edu), Bruce Cornuelle, and William A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., Univ. of California, San Diego, San Diego, CA)

The waveguide invariant (WGI) is a property that can be used to localize acoustic radiators and extract information about the environment. Here, the WGI is determined using ships as sources of opportunity, tracked using the Automatic Identification System (AIS). Using a single hydrophone the acoustic intensity as a function of range and angle is measured in the presence of ships at a variety of ranges and angles. The relationship between range, intensity, and frequency is used to determine the WGI parameter β for ships in a variety of positions. These Beta values are interpolated and a map of β is created for the environment in question. The method is demonstrated using data collected on a single hydrophone in a shallow water environment off the coast of Southern California.

Session 5aBAa

Biomedical Acoustics: Acoustics for Older Persons with Disabilities

Takayuki Arai, Cochair
Information and Communication Sciences, Sophia University, 7-1 Kioi-cho, Chiyoda-ku, Tokyo 102-8554, Japan

Mari Ueda, Cochair
Department of Design Strategy, Kyushu University, 4-9-1 Shiobaru, Fukuoka 815-8540, Japan

Invited Papers

8:00

5aBAa1. The activities of the research committee on "Oto barrier-free" of the Acoustical Society of Japan. Kimio Shiraishi (Faculty of Design, Kyushu Univ., 4-9-1 Shiobaru, Minami-ku, Fukuoka 815-8540, Japan, kimio@design.kyushu-u.ac.jp), Kentaro Nakamura (Inst. of Innovative Res., Tokyo Inst. of Technol., Tokyo, Japan), Yasuhiro Oikawa (Faculty of Sci. and Eng., Waseda Univ., Tokyo, Japan), and Mari Ueda (Faculty of Design, Kyushu Univ., Fukuoka, Japan)

The number of the persons above 65 years old has reached 26% of the population in Japan. Consequently, the number of persons with hearing and/or visual impairments has increased very rapidly. It is important to support communication or sound information compensation for these persons. From this background, the research committee on "Oto barrier-free" of the Acoustical Society of Japan has started in 2006. The "Oto barrier-free" committee intends to resolve problems related to improving accessibility in daily life for persons with disabilities or elderly people. The research fields are very diverse and include speech communication, electro- acoustics, noise and vibration, architectural acoustics, and laws for disabled persons. The research activities for persons with hearing impairments or elderly people aim to improve their listening environment in indoor spaces, such as homes for the aged or station buildings, by testing the use of sound absorbing materials and hearing assistive devices. The research assists or complements the use of aids such as voice guidance devices for persons with a visual impairment, and will be useful for foreigners visiting the Tokyo Olympics and Paralympics Games in 2020 too.

8:20

5aBAa2. Relationship between auditory degradation and fricatives/affricates production and perception in elderly listeners. Keiichi Yasu (National Rehabilitation Ctr. for Persons with Disabilities, 1, Namiki 4-chome, Tokorozawa 359-8555, Japan, keiichi.yasu@gmail.com), Takayuki Arai (Sophia Univ., Tokyo, Japan), Kei Kobayashi (The Univ. of Auckland, Auckland, New Zealand), and Mitsuko Shindo (Sophia Univ., Tokyo, Japan)

Elderly people often complain that they struggle with consonant identification when listening to spoken language. In previous studies, we conducted several experiments, including identification tests for young and elderly listeners using /shi/-/chi/ (CV) and /ishi/-/ichi/ (VCV) continua. We also recorded production of such CV and VCV syllables. For the CV stimuli, confusion of /shi/ as /chi/ increased when the frication had a long rise time. The degree of confusion increased when auditory property degradation was observed such as threshold elevation in high frequency. In the VCV stimuli, confusion of /ichi/ as /ishi/ occurred for a long silent interval between the first V and C with auditory property degradation. In the present study, we analyzed the utterances of those CV and VCV syllables and measured the duration of frication and silent interval. The direction of the boundary shift in the perception of fricatives and affricates by auditory degradation was consistent with that of production. We found that degradation of auditory properties affects both perception and production of fricatives and affricates in elderly people.

8:40

5aBAa3. A speech synthesizer of Japanese saves patients' voices. Takayuki Arai (Sophia Univ., 7-1 Kioi-cho, Chiyoda-ku, Tokyo 102-8554, Japan, arai@sophia.ac.jp), Shigeto Kawahara (The Keio Inst. of Culture and Linguistic Studies, Tokyo, Japan), Musashi Homma (Tokyo Metropolitan Neurological Hospital, Tokyo, Japan), and Takaki Yoshimura (Pasobora, Sasebo, Japan)

"MyVoice," free speech synthesis software for Japanese, developed by one of the authors, is widely used to save many patients' voices in Japan. We lose our voice for various reasons. Amyotrophic Lateral Sclerosis, or ALS, is an example. When an ALS patient has difficulty breathing, Tracheotomy becomes an option, and the patient loses the ability to speak. Laryngectomy is another example. By using "MyVoice," patients can keep communicating using their own voice even after losing that ability. To facilitate this, patients record utterances with their voice before surgery. The recording is designed to minimize time and load of patients as much as possible, by including Japanese basic mora set so that potentially any word can be concatenated from the recorded sounds. The graphical user interface of this software is well-designed so that therapists as well as the patients' family and friends can also use it with a minimum of effort. In fact, we often hear heart-warming stories of making "MyVoice." "MyVoice" continues to rescue patients' voices and we will continue making improvements to it in the future.

5a FRI. AM

3365   J. Acoust. Soc. Am., Vol. 140, No. 4, Pt. 2, October 2016

5th Joint Meeting ASA/ASJ   3365

9:00

5aBAa4. Audio service for elderly people in broadcasting. Tomoyasu Komori, Takehiro Sugimoro, and Atsushi Imai (Sci. & Technol. Res. Labs., NHK (Japan broadcasting corporation), 1-10-11 Kinuta Setagaya-ku, Tokyo 1578510, Japan, komori.t-bw@nhk.or.jp)

Elderly viewers sometimes feel that the background sound (music and sound effects) in television programs is too loud or that the narration or speech is too fast to understand. We attempted to solve these problems by two methods: improvement of the receiver system and the development of a new broadcasting system. We prototyped a receiver system in which loudness balance between the dialog (narration or speech) and background sound can be adjusted, as well as the speech rate. The results of subjective evaluation experiments with elderly viewers showed that the use of this system could significantly facilitate listening to television sound. In 2016, we started test broadcasting of 8K Super Hi-Vision (8K SHV) programs that have highly immersive surround sound produced through 22.2 multi-channel (22.2ch). With the aim of improving the functionality of 8K SHV sound as a broadcasting service, a domestic standard has been established for the dialogue control functions, including the dialog enhancement and a dialog replacement function. In this presentation, we show these two types of audio service for elderly people.

9:20

5aBAa5. Use of data hiding in aerial sounds for accessibility. Akira Nishimura (Dept. of Informatics, Tokyo Univ. of Information Sci., 4-1, Onaridai, Wakaba-ku, Chiba 2658501, Japan, akira@rsch.tuis.ac.jp)

This article reviews previous technologies for aerial acoustic data transmission, namely, audio data hiding, acoustic modems, and a hybrid of the two. These technologies can provide an information channel to convey index data for static text messages or data for text messages to assist people who are deaf or hard of hearing. The payload bitrate of these technologies ranges from several bits per second for long distances (>100 m) to hundreds of bits per second for short distances (<1 m). Previous studies have mainly focused on developing methods for embedding and encoding information into aerial sounds. However, few works have evaluated actual aerial transmission in the presence of various disturbances, such as reverberation and reflections, transfer functions of loudspeakers and microphones, non-linear distortions in loudspeakers, background noise, and Doppler effects. Therefore, in this study, evaluation methods for these acoustic data transmission technologies are proposed and discussed. In addition, practical applications are discussed regarding the use of the technologies as a means of accessibility for people with speech and language disabilities.

9:40

5aBAa6. AcousessMap: Smartphone-based collaborative tool to facilitate assessing acoustical accessibility conditions for visually impaired people. Mari Ueda (Dept. of Design Strategy, Kyushu Univ., 4-9-1 Shiobaru, Fukuoka, Minami-ku 815-8540, Japan, m-ueda@design.kyushu-u.ac.jp), Takahiro Miura, Ken-ichiro Yabu (Univ. Tokyo, Tokyo, Japan), Takashi Morihara (INCT, Kanazawa, Japan), and Yoshio Tsuchida (K I T, Kanazawa, Japan)

Assistive instruments such as textured paving blocks and acoustical traffic signals are installed for helping visually impaired people to walk comfortably and avoid traffic accidents, who have any inconvenience to move outside. Though these precipitous situations of accessibility progress affect their migration pathway for their destination, up-to-date accessibility information is difficult to gain quickly because of local information disclosure. Thus it is necessary to develop a comprehensive system that appropriately acquires and arranges scattered accessibility information, and then presents this information intuitively. However, these systems present volunteers with difficulties when they are gathering accessibility conditions and then organizing them. Also, most of the volunteers do not know the current situations of acoustical support systems for the visually impaired. In this study, our final goal is to establish an efficient scheme to share the conditions and places of the acoustical support systems by many volunteers. Particularly, in this report, we aimed to check the feasibility to share the acoustical conditions to the web by persons without any knowledge of the support systems. We first developed a smartphone-based application for sharing accessibility conditions and carried out events to share accessibility conditions as a part of a lecture course for technical college students.

Session 5aBAb

Biomedical Acoustics: Biomedical Acoustics Poster Session

Kang Kim, Chair
Medicine, University of Pittsburgh, 950 Scaife Hall, 3550 Terrace Street, Pittsburgh, PA 15261

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

Contributed Papers

5aBAb1. Tissue boundary roughness index for evaluating malignancy of cancer in ultrasound-computed tomography system. Kazuhiro Yamanaka, Takahide Terada, Yushi Tsubota, Wenjing Wu, Atsuro Suzuki, Takashi Maruoka, and Ken-ichi Kawabata (Res. & Development Group, Hitachi, Ltd., 1-280, Higashi-Koigakubo, Kokubinji, Tokyo 185-8601, Japan, kazuhiro.yamanaka.uu@hitachi.com)

Ultrasound-computed tomography (USCT) is a promising candidate for a radiation free, painless, and quantitative modality for breast cancer examination. We developed a USCT prototype with a ring-shaped transducer array (frequency: 1.5 MHz) that moves along the ring axis for scanning the entire breast. Distributions of speed and attenuation of ultrasound are computed from ultrasound transmitted through tissue. Although prior research shows that there is a correlation between malignancy and the speed of sound and attenuation values, in some cases both benign and malignant tumors have similar values. In this study, we propose to quantify the boundary roughness of a tumor from the spatial power distributions of reflected ultrasounds that are transmitted from different apertures of the ring transducer. In conventional ultrasound echography, the boundary roughness of a tumor is an important indicator of malignancy. However, roughness is judged by radiologists based on B-mode images and is a subjective index. We simulated the ultrasound reflection signal from boundaries of different roughness. The simulation result indicates that a spatial power concentration has a significant correlation with boundary roughness. We confirmed through gel phantom experiments that the index distinguishes rough and smooth boundaries, which is not evident in B-mode images.

5aBAb2. Tissue lesion formation induced by high intensity focused ultrasound along a spiral pathway. Dong Zhang, Juan Tu, Xiasheng Guo (Phys., Inst. of Acoust., Nanjing Univ, 22 Hankou Rd., Nanjing 210093, China, dzhang@nju.edu.cn), Qingyu Ma, and Gepu Guo (Phys., Nanjing Normal Univ, Nanjing, China)

Both theoretical and experimental studies were performed here to investigate the lesion formation induced by high-intensity focused ultrasound (HIFU) operating in continuous scanning mode along a spiral pathway. The Khokhlov-Zabolotskaya-Kuznetsov equation and bio-heat equation were combined in the current model to predict HIFU-induced temperature distribution and lesion formation. The shape of lesion and treatment efficiency were assessed for a given scanning speed at two different grid spacing (3 mm and 4 mm) in both the gel phantom and *ex vivo* studies. The results show that uniform lesions can be generated by the homogenization of thermal diffusion along the spiral scanning pathway. The complete coverage of the entire treated volume can be achieved as long as the spacing grid of the spiral pathway is smaller than a critical value that right matches the maximum thermal diffusion dimension, and the treatment efficiency can be optimized by selecting an appropriate scanning speed. This study can provide guidance for further optimization of the treatment efficiency of HIFU therapy.

5aBAb3. Relative impedance variation based temperature monitoring and dose control for high intensity focused ultrasound therapy. Qingyu Ma, Gepu Guo (School of Phys. andTechnol., Nanjing Normal Univ., 1 Wenyuan Rd., Nanjing, Jiangsu 210023, China, maqingyu@njnu.edu.cn), Juan Tu, Dong Zhang (Inst. of Acoust., Nanjing Univ., Nanjing, Jiangsu, China), and Shaotong Feng (School of Phys. andTechnol., Nanjing Normal Univ., Nanjing, Jiangsu, China)

High intensity focused ultrasound (HIFU) can be used as a non-invasive and efficient therapy approach to treat tumors by heat coagulation with less metastasis, although it is still limited by inaccurate temperature monitoring and efficacy evaluation. Based on the relationship between temperature and electrical impedance, a model of impedance measurement during HIFU therapy was established. Numerical studies on the distributions of acoustic pressure, temperature and conductivity were simulated, and the impedances in the model were calculated with the electrical potential and current density. It is proved that, for a fixed acoustic power in the focal region, the relative impedance variation (RIV) linearly increases with the increasing treatment time and is inversely proportional to acoustic power, while the relative impedance variation rate (RIVR) shows a linear relationship with the acoustic power. The treatment time that is required for the temperature in the focal region to reach 70 degree is inversely proportional to the square of acoustic power. The experimental results of RIV and RIVR measured at different powers and treatment times are in good agreements with the simulated data. The results suggest that the RIV can be used as a useful indicator for the non-invasive temperature monitoring and efficacy evaluation, and might provide new strategy for accurate dose control of HIFU therapy.

5aBAb4. A high-frequency ultrasound method for testing skin cancer margins ex vivo. Zachary A. Coffman (Biology, Utah Valley Univ., 2750 W. Spaulding Ln., West Jordan, UT 84088, zachary.a.coffman@gmail.com), Caitlin Carter (BioTechnol., Utah Valley Univ., Orem, UT), Dolly A. Sanjinez (Biology, Utah Valley Univ., Orem, UT), Garrett Wagner (Comput. Eng., Utah Valley Univ., Orem, UT), Robyn K. Omer (Botany, Utah Valley Univ., Orem, UT), and Timothy E. Doyle (Phys., Utah Valley Univ., Orem, UT)

Mohs surgery is the standard surgical treatment for basal cell carcinoma and several melanomas. It involves resecting successive layers of the skin cancer and surrounding tissue (margins), analyzing the tissue in a pathology laboratory, and using the results to determine whether all of the cancer has been removed or whether another layer needs to be resected. Although resection of each layer only takes 5-10 minutes, procedure times for Mohs surgery are approximately 4 hours since the pathology analysis is very time consuming. A rapid specimen assessment technique for detecting skin cancer in surgical margins would therefore be of significant benefit to dermatology surgeons and their patients. A high-frequency ultrasound method has been developed to test Mohs surgical specimens without the need for direct immersion or transducer contact on the specimens. The ultrasound system

5a FRI. AM

uses a 50-MHz, narrow-beam (1.5 mm) pachymeter probe; a 50-MHz, broad-beam (6.25 mm) transducer; a high-frequency pulser-receiver; a 1-GHz digital oscilloscope; and an aluminum test stage to hold and scan the specimen. Pitch-catch and pulse-echo measurements are collected from each test position on the specimen to determine the position's thickness and ultrasonic wave speed. Other parameters calculated from the data include attenuation and spectral peak density.

5aBAb5. Transducer influence on magnetoacoustic tomography with magnetic induction based on 3D equivalent source analysis for tissues. Qingyu Ma, Shaotong Feng, Gepu Guo (School of Phys. andTechnol., Nanjing Normal Univ., 1 Wenyuan Rd., Qixia District, Nanjing, Jiangsu 210023, China, maqingyu@njnu.edu.cn), Juan Tu, and Dong Zhang (Inst. of Acoust., Nanjing Univ., Nanjing, Jiangsu, China)

As a noninvasive imaging modality based on the coupling of magnetic and acoustic fields, magnetoacoustic tomography with magnetic induction (MAT-MI) has been demonstrated to have the capability of imaging the variation of conductivity distribution inside the object. However, the image resolution is still limited by the parameters of the receiver. Based on the theory of acoustic dipole, 3D equivalent source analysis was used to simulate transducer detected pressures and waveforms for MAT-MI. The influence of transducer was studied both theoretically and experimentally for a cylindrical model. It is demonstrated that large-radius transducer with strong reception pattern can detect the acoustic signals transmitted along its normal direction with sharp pressure peaks, reflecting the divergence of the induced Lorentz force. By considering the effect of acoustic attenuation and the accuracy of image reconstruction, the acoustic pressure with acceptable peak pressure ratio and improved SNR can be detected when the scanning distance is 5-10 times the radius of the object. Wide bandwidth transducer should also be selected to reduce the boundary width of borderline stripes and improve the spatial resolution of reconstructed images. The favorable results confirm the influence of transducer on MAT-MI and also provide the fundamentals for transducer selection in further study to improve the accuracy of electrical impedance reconstruction.

5aBAb6. Optimization of sphingosylphosphorylcholine concentration for detection of cytoskeletal changes in malignant pancreatic cells using high-frequency ultrasound. Caitlin Carter (BioTechnol., Utah Valley Univ., 886 E. Old English Rd., Draper, UT 84020, caitlin.carter03@gmail.com), Dolly Sanjinez (Biology, Utah Valley Univ., Orem, UT), Mandy Marvel (BioTechnol., Utah Valley Univ., Orem, UT), and Timothy E. Doyle (Phys., Utah Valley Univ., Orem, UT)

The biomechanical properties of cells can be greatly influenced by modifications of the cytoskeleton of the cell. Cytoskeletal modifications can often be linked to disease and disease pathways. Cancer metastasis has been previously associated with the cytoskeletal modifications that are induced by the addition of sphingosylphosphorylcholine (SPC) to panc-1 cells. SPC is a bioactive lipid that reorganizes keratin filaments in the cytoskeleton into a perinuclear shape that enables cell mobility. The maximum effect of SPC on the keratin reorganization has been previously observed as being concentration dependent, with $5\text{-}10\mu M$ concentration being optimal as observed through keratin immunostaining. In this work, high frequency ultrasound was used to observe the SPC induced keratin reorganization through an ultrasound signature. Ultrasound measurements were obtained using $10\mu M$ SPC and $\sim 17\mu M$ SPC concentrations. Significant time-dependent changes in the ultrasound measurements were observed in the higher concentration of SPC as compared to control cell cultures that were not treated with SPC. In contrast, the lower concentration of SPC did not show visible changes in the ultrasound measurements taken. Future work will include optimizing a more specific concentration of SPC for observing changes in the ultrasound measurements taken of panc-1 cells.

5aBAb7. Modular arrays for transpinal ultrasound application. Shan Qiao, Constantin Coussios, and Robin Cleveland (Dept. of Eng. Sci., University of Oxford, Inst. of Biomedical Eng., Old Rd. Campus Res. Bldg., Headington, Oxford, Oxford OX3 7DQ, United Kingdom, shan.qiao03@gmail.com)

Lower back pain is one of the most common health problems in developed countries, and is normally caused by the degeneration of intervertebral discs. High intensity focused ultrasound can be used to mechanically fractionate degenerate disc tissue by inertial cavitation. Due to the complexity of the spine structure, delivering sufficient focused acoustic energy to the target zone without damaging surrounding tissue is challenging and further exacerbated by patient-to-patient variability. Here we propose the design of modular arrays, each consisting of 32 elements at 0.5 MHz, which can be configured to optimize delivery for a specific patient by the means of time-reversal using the patient geometry derived from CT scans. Initial tests were carried out in water and compared with simulations. Using four different configurations of four modules, focusing was accomplished using a needle hydrophone at the target location to determine the transmit phase for each element. In each of the four configurations a focal gain of approximately 35 was achieved and steering range of the focus was +/-30 mm in azimuth, +/-5mm in elevation. These results confirm that the numerical model accurately predicts the acoustic field of the modular array and can be used for treatment planning in a disc. [Work supported by EPSRC, UK.]

5aBAb8. Dynamics of an acoustically driven liquid-gas interface. Brandon Patterson and Eric Johnsen (Univ. of Michigan, 1231 Beal Ave., Rm. 2016, Ann Arbor, MI 48109, awesome@umich.edu)

Interactions between acoustic waves and liquid-gas interfaces occur in a variety of applications, including diagnostic ultrasound of the lung. We hypothesize that, because of the sharp density gradient at the fluid discontinuity, acoustic waves may generate sufficient baroclinic vorticity to appreciably deform liquid-gas interfaces. This nonlinear effect has been studied for shock-accelerated interfaces, i.e., the Rictmyer-Meshkov instability; however, it is not describable through traditional, linear acoustic methods and is not well studied for acoustic waves. To investigate our hypothesis, we simulate an inviscid, compressible fluid system with acoustic waves impinging from water into air. We find that acoustic waves capable of deforming the interface during the wave-interface interaction period will deposit lasting baroclinic vorticity along the interface, capable of driving deformation that continues long after the passage of the wave. We use dimensional analysis to describe the dependence of the late-time interface growth on the deposited vorticity. Lastly, we show results to demonstrate that for acoustic waves with properties relevant to diagnostic ultrasound encountering a water-air interface, nearly all of the vorticity is generated in gas-dominated fluid. We explain this using an order of magnitude analysis.

5aBAb9. Parametric array for tissue harmonic imaging. Christina P. Keravnou and Michalakis A. Averkiou (BioEng., Univ. of Washington, William H. Foege Bldg., 3720 15th Ave. NE, Seattle, WA 98195, maverk@uw.edu)

More than 50% of abdominal scans are performed on technically difficult patients due to worldwide increase of obesity. Tissue harmonic imaging has shown some improvements on obese patients but the problem remains unsolved. The use of dual frequency pulses allows the generation of sum and difference frequencies in tissue. We hypothesize that the difference frequency (Δf, parametric array) can survive larger penetration depths and increase SNR while maintaining the spatial characteristics of the primary wave. Our objective was to optimize the parameters for efficient generation of the parametric array and study its spatial characteristics for imaging obese patients. Numerical simulations using KZK equation and measurements in tissue-like media were performed to study parametric arrays. Pulses with two frequencies (f_1, f_2) were considered where $0.1f_1 \leq \Delta f \leq f_1$. Pulse inversion was utilized to isolate even harmonics, including Δf. The amplitude of Δf was maximized at $\Delta f = f_1$. The axial resolution is not seriously compromised by the lower frequency of Δf when using PI. For 0.5 dB/cm attenuation difference between Δf and $2f1$, Δf echoes from 20 cm deep were greater by 10 dB. Application of the parametric array on larger technically difficult patients would offer similar benefits as those in sonar applications.

5aBAb10. Vascular vector flow mapping for ultrasonic two-dimensional flow estimation. Rei Asami, Tomohiko Tanaka, Motochika Shimizu, Yoshinori Seki, Tomohide Nishiyama, Hajime Sakashita, and Takashi Okada (Hitachi, Ltd., 1-280 HIgashi Koigakubo, Kokubunji, Tokyo 185-0003, Japan, rei.asami.fq@hitachi.com)

Vector flow mapping (VFM) is an echocardiographic approach for visualizing two-dimensional cardiac flow dynamics by estimating the radial component of flow from the Doppler velocities and wall motion velocities using mass-conservation equations. Although VFM provides two-dimensional flow in a clinically suitable, fast, and easy manner, it is only applicable to bounded regions such as ventricles. In this study, the VFM algorithm is modified so that the velocities are estimated regardless of the flow geometry. The proposed method, vascular VFM, operates on the same principle as VFM but with additional boundary conditions created by using Doppler velocities measured at a different steering angle. The method was optimized and validated using a common carotid artery phantom and particle-image-velocimetry (PIV). The results indicated that the optimal angle of an additional Doppler beam ranged from 10 to 20 degrees. Under this optimal-beam-angle condition, the standard deviation (SD) of total vascular VFM error, normalized by maximum velocity, is reasonably small (8.1 to 16.3%). The error is predominantly from input Doppler velocities, which has error of up to 12.7%. When calculated with ideal inputs, the algorithm error ranges from 1.9 to 4.3%. These results indicate that vascular VFM algorithm estimates two-dimensional flow accurately in the non-bound region and may be clinically significant in cases such as arteriosclerosis diagnosis.

5aBAb11. Quantitative evaluation of hemolysis induced by pulsed ultrasound exposure. Kentaro Tani, Masato Imura (Graduate of Sci. and Eng., Doshisha Univ., 1-3 Miyakodani,Tatara610-0349, Kyotanabe, Kyotanabe-shi 610-0349, Japan, duq0369@mail4.doshisha.ac.jp), Daisuke Koyama (Faculty of Sci. and Eng., Doshisha Univ., Kyotanabe, Japan), and Yoshiaki Watanabe (Faculty of Life and Medical Sci., Doshisha Univ., Kyotanabe, Japan)

To establish safety criteria of ultrasound for blood, we examined hemolysis induced by low-intensity pulsed ultrasound and performed *in vitro* experiments using bovine blood. Quantitative evaluation by microscopic observations and the relationship between the hemolysis and cavitation microbubbles were discussed. Hemolysis was evaluated by red blood ghost cell. The presence of red blood ghost cells in the plasma component implies that the cell membrane was broken and the internal hemoglobin flowed out to surrounding blood plasma. Pulsed ultrasound at 1 MHz with the maximum sound pressure amplitude of 200 kPa was irradiated for one minute, and the ratio of non-exposure time (0—80%) was changed. After ultrasound exposure, the blood samples were centrifuged and divided into two layers: blood cells and plasma components. 10 μL of the plasma components was sampled and the number of the red blood ghost cells was counted. The number of red blood ghost cells increased gradually with the decrease in the ratio of non-exposure time and decreased again at 0% (continuation wave). This result implies that the hemolysis is mainly caused by microbubble cavitation and the non-exposure time of ultrasound is one of the important factors for hemolysis in such low-intensity ultrasound field.

5aBAb12. Controlled nanoparticle release through microbubble microstreaming: Physical insight and applications. Guillaume Lajoinie (Phys. of Fluids Group, Univ. of Twente, P.O. Box 217, Enschede 7500 AE, Netherlands), Ying Luan (Biomedical Eng., Erasmus MC, Rotterdam, Netherlands), Erik C. Gelderblom (Phys. of Fluids Group, Univ. of Twente, Enschede, Netherlands), Benjamin Dollet (Institut de Physique de Rennes, Rennes, France), Heleen Dewitte, Ine Lentacker (Faculty of Pharmaceutical Sci., Ghent Res. Group on Nanomedicines, Ghent, Belgium), and Michel Versluis (Phys. of Fluids Group, Univ. of Twente, Enschede, Netherlands, m.versluis@utwente.nl)

Microbubbles driven by ultrasound are used in a number of applications including surface cleaning, ultrasound imaging, and as a vehicle for local drug delivery. To prolong the microbubble lifetime, its gas core is coated with a stabilizing shell, typically consisting of phospholipids. The coating can also be used to attach a payload of functional nanoparticles. Interestingly, upon ultrasound irradiation at several hundreds of kPa, the payload was observed to be released in a highly controlled way. This release carries great potential for using microbubbles as drug delivery agents in the context of personalized medical therapy. However, until now, limited experimental observations of the phenomenon are available. Here, we study using ultra high-speed and fluorescence imaging techniques in top and in side-view the underlying mechanisms of the release. We also developed a model on the basis of a Rayleigh-Plesset-type equation that reveals that non-spherical bubble behavior is key to the release mechanism. We also quantified the streamlines and acoustic streaming velocity responsible for the microfluidic transport of the material and propose for the first time a complete physical description of the controlled release of ultrasound actuated microbubbles.

5aBAb13. All-in one small modular transducer system for ultrasound therapy. Kiyoshi Yoshinaka, Hideki Takeuchi, Shingo Toyoda, Akira Sasaki (Health Res. Inst., AIST, AIST East-Build, 1-2-1, Namiki, Tsukuba, Ibaraki 3058564, Japan, k.yoshinaka@aist.go.jp), Takashi Azuma, and Shu Takagi (Univ. of Tokyo, Tokyo, Japan)

We developed new concept all-in one small modular transducer system for ultrasound therapy. Focus position control of HIFU (High Intensity Focused Ultrasound) with multi-elements phase delay control is very popular in clinical application. However, multi-elements driving amplifiers are very large size like a large refrigerator. So, we adopted a direct drive amplifier system for the multi-elements modular transducer. The modular system has some advantages compared to conventional systems. It can be reduced the energy loss of the connecting cable between the transducer element and the amplifier. Other advantages are below. Free placement of module transducer elements like a patient specific body line. Easy maintenance of each element. Easy design change to reduce the manufacturing time. The modular transducer system is composed of four blocks. The first module is multi-channel transducer elements. The second module is multi-channel direct drive amplifier system. The third module is a name-card size multi-channel phase delay signal generator with the phase delay signal set memory bank. The fourth module is a phase delay signal set generating software with a personal computer. It can transmit the signal set to the third module. We manufactured the prototype very small all-in one modular transducer system for ultrasound therapy.

5aBAb14. Detection of tumoral netrin-1 protein levels using ultrasound microbubble contrast agents: *In vitro* validation. Jennifer Wischhusen (LabTAU IINSERM 1032, French National Inst. of Health and Medical Res. (Inserm) & Université de Lyon, 151 Cours Albert Thomas, Lyon 69390, France), Jean-Guy Delcros, Patrick Mehlen (Inserm U1032, Res. Cancer Ctr. of Lyon, Lyon, France), and Frederic Padilla (LabTAU IINSERM 1032, French National Inst. of Health and Medical Res. (Inserm) & Université de Lyon, Lyon, France, frederic.padilla@inserm.fr)

Netrin-1 is a protein secreted by tumor cells in breast cancers. It triggers survival and proliferation through its interaction with the dependence receptor UNC5B. Netrin-1 is the target for a therapy under development. To identify patients who can benefit from treatment, a companion diagnostic is required. Our aim is to characterize netrin-1 as a target for ultrasound molecular imaging and to validate netrin-1-targeted microbubbles. Netrin-1 interactions with the extracellular matrix were analyzed by flow cytometry and Western blot. Microbubbles were functionalized and binding under static and dynamic conditions on purified netrin-1 or netrin-1-expressing cells was assessed by light microscopy. Netrin-1 interacted with heparin sulfate proteoglycans on the surface of the cells and with the cell surface receptor UNC5B. Microbubbles specifically and dose-dependently bound to netrin-1 protein and netrin-1-expressing cells, in absence and presence of UNC5B. Our findings confirmed that netrin-1, although being a secreted protein, is a potential target for ultrasound molecular imaging using microbubbles that specifically attach to netrin-1 presented on the cell surface. Currently, netrin-1 presentation on the luminal side of the vascular endothelium and ultrasound molecular imaging in the living animal are assessed.

5a FRI. AM

5aBAb15. Surgeon assessment versus laboratory measurement: brain tissue mechanics quantified by expert opinion. Sheronica L. James (Chemical and Biomedical Eng., Cleveland State Univ., 2121 Euclid Ave., Cleveland, OH 44115, sheronicajames@gmail.com), Mark Howell, Qi Wang, and Gregory T. Clement (Biomedical Eng., Cleveland Clinic, Cleveland, OH)

Transcranial ultrasound (tUS) is a plausible means of measuring motion of the brain due to traumatic force or structural abnormalities. However, limited data in the literature on the mechanical properties of human brain tissue make it practically impossible to develop a mechanically realistic ultrasound phantom that can mimic the brain under these conditions. The present study aimed to determine, by indirect qualitative assessment, the stiffness of in-vivo brain tissue. Six surgeon examiners under blinded conditions physically examined 16 gelatin-based tissue phantoms (0.020-0.095 g/mL) to determine which sample best matched the stiffness of brain tissue in-vivo. Ten cylindrical phantom samples, acoustically matched to human brain tissue, were prepared based on the examiners' responses (0.030±0.008 g/mL), and stress-strain behavior was characterized under compression up to 50% strain. The average compressive stress at 50% strain was 10.3±1.4 kPa. These results correlate well with those previously reported in the literature, and add a practical measure of brain stiffness to the paucity of information available on human brain tissue mechanics. Furthermore, it is potentially an initial step in developing mechanically realistic models of structural disorders of the brain and traumatic brain injury to evaluate theories on the brain's mechanical behavior under these circumstances.

5aBAb16. Bio-acoustic levitational assembly of heterocellular multilayer constructs for tissue engineering. Charlene Bouyer (LabTAU INSERM 1032, French National Inst. of Health and Medical Res. (Inserm) & Université de Lyon, 151 Cours Albert Thomas, Lyon 69390, France), Pu Chen, Utkan Demirci (Bio-Acoust. MEMS in Medicine (BAMM) Lab, School of Medicine, Stanford Univ., Stanford, CA), and Frederic Padilla (LabTAU INSERM 1032, French National Inst. of Health and Medical Res. (Inserm) & Université de Lyon, Lyon, France, frederic.padilla@inserm.fr)

Many tissues are comprised of multiple heterocellular layers, where interlayer cell communications between heterogeneous cell populations are essential to sustain normal tissue functions. Mimicking such natural organization presents a challenge to tissue engineering, and few methods are available to bioengineer heterocellular multilayer constructs in vitro. We have developed a bulk acoustic levitation (BAL) technique to assemble multilayer 3D neuronal constructs from homogenous neural progenitor [1]. We present here a novel approach to build heterocellular multilayer tissues by combining BAL and layer-by-layer assembly. Cells are levitated into several layers within fibrin prepolymer solution by bulk acoustic waves and immobilized in a fibrin hydrogel layer via gelation. Several hydrogel units can then be stacked subsequently, each unit with a distinct cell type, to form a single hydrogel construct. We demonstrate the bioengineering of a heterogenous multilayer construct composed of three different cell types (wild-type—mCherry and eGFP HeLa cells) spatially distributed in a single 3D hydrogel. This novel approach enables rapid formation 3D multilayer structures in minutes with tunable interlayer spacing, layer composition and thickness, and is potentially useful to bioengineer homogenous and heterogeneous multilayer cell organization for tissue engineering. [1] Bouyer et al. (2015) Adv. Mater. **28**:161-7

5aBAb17. Locally-concentrated sonicated microbubbles: Potential as an antibiotic. Mark Howell, Qi Wang (Biomedical Eng., Cleveland Clinic, 9500 Euclid Ave., ND20, Cleveland, OH 44195, howellm2@ccf.org), Shota Shimizu (Eng., Mie Univ., Ikoma-gun, Nara-ken, Japan), Albert Feeny (Lerner College of Medicine, Cleveland Clinic, Cleveland, OH), and Greg Clement (Biomedical Eng., Cleveland Clinic, Cleveland, OH)

While the antimicrobial effects of ultrasound have been well established, application as a bactericide in medicine has been hindered in part by significant variation in reported efficacy. Non-thermal mechanisms of cell death are generally attributed to cavitation, prompting us to investigate the potential for locally applied microbubbles to serve as a catalyst for reliably increasing bacterial death, while sparing surrounding cells. This preliminary study was limited to gram-negative bacteria (E-coli) suspended in buffered saline. Preliminary work at 40 kHz continuous-wave sonication established a threshold time for our setup of 15 minutes, below which negligible cell death was observed relative to control samples. Experiments were then performed with and without the introduction of octofluropropane-filled lipid shells microbubbles, sonicated over the 15 minute period. Thermometry and passive cavitation monitoring (PCD) were both conducted. Analysis by dilution and plating indicate a 65-80% reduction in bacteria relative to controls, while microscopy, cavitation detection, and a negligible temperature rise all point to cavitation as the mechanism of death. Results motivate further investigation into the utility of microbubbles in combination with ultrasound for the treatment of certain non-perfusing infections such as chronic wounds and abscesses. Further work is necessary to confirm similar effects for gram-positive bacteria.

5aBAb18. Numerical investigation of the nonlinear dynamics of interacting microbubbles. Amin Jafari Sojahrood, Hossein Haghi, Raffi Karshafian, and Michael C. Kolios (Dept. of Phys., Ryerson Univ., 350 Victoria St., Toronto, ON M5B2K3, Canada, amin.jafarisojahrood@ryerson.ca)

The successful application of MBs requires detailed understanding of the nonlinear behavior of microbubbles (MBs), especially in polydisperse clouds, where the dynamic of every individual MB affects the other MBs. However, there is not enough information on how the nonlinear oscillations of one MB influences the other and vice versa. In this work the dynamics of 2 and 3 interacting MBs of initial radii of $1\mu m < R0 < 4$ μm are studied by investigating the pressure dependent resonance curves and the bifurcation diagrams of the MBs (1 MHz$<$f$<$15 MHz, 1 kPa$<$p$<$1 MPa) for cases of no interaction and interaction with varying MB-MB distances. Results show that, for small enough distances, the pressure dependent resonance frequencies (fr) and the pressure threshold for harmonic and SH fr decreases. The larger MB may force new peaks in the resonance curves of the smaller MB at fr, harmonic fr and SH fr of the larger MB. The larger MB can control the dynamics of the smaller MB and force the smaller MB to exhibit the same nonlinear behavior (e.g., $1/2$, 1/3, 1/4 SH) as of the larger MB. The dynamics of a cloud of interacting MBs maybe controlled by controlling the dynamics of the larger MB in the population.

5aBAb19. High-throughput production of microbubble contrast agents using a sonofluidic device. Eleanor P. Stride, Richard Browning, Paul Rademeyer, and Dario Carugo (Univ. of Oxford, Old Rd. Campus Res. Bldg., Oxford OX3 7DQ, United Kingdom, eleanor.stride@eng.ox.ac.uk)

The response of microbubbles to a given sound field is determined by their size and coating. These, in turn, depend on their chemical formulation and the production technique. Sonication is the most commonly employed method and can generate high concentrations of microbubbles rapidly but with a broad size distribution and poor reproducibility. Microfluidic devices provide excellent control over size, but the small-scale architectures required are often challenging to manufacture, offer low production rates, and are prone to clogging. Microbubbles may also have inferior surface characteristics and stability compared to those produced by sonication. In this study we investigate a hybrid technique in which monodisperse microbubbles of ~200μm diameter are produced at high flow rates in a simple T-junction and then undergo controlled fragmentation by exposure to ultrasound via an integrated transducer operating between 71-73kHz. Microbubbles were prepared using the device or a standard sonication protocol and compared in terms of their size, size distribution, concentration, stability, acoustic response, and surface molecular concentration using quantitative fluorescence microscopy. The characteristics of the microbubbles produced by the device were found to be equivalent in terms of production rate, stability and acoustic response but with a narrower size distribution and tunable mean size.

5aBAb20. Propagation of low amplitude focused ultrasound waves in a bubbly medium: Finite element simulation and theoretical considerations for optimization of the treatment parameters. Amin Jafari Sojahrood, Hossein Haghi, Raffi Karshafian, and Michael C. Kolios (Dept. of Phys., Ryerson Univ., 350 Victoria St., Toronto, ON M5B2K3, Canada, amin.jafarisojahrood@ryerson.ca)

Administration of microbubbles (MBs) increases the attenuation and changes the sound speed (Cs) of the medium. These phenomena are nonlinear and depend on MB size frequency and acoustic pressure. The increased attenuation can shield the focal region and reduce the efficacy of the application (e.g., drug delivery). Detailed information on the attenuation and Cs in the medium is needed to optimize these applications. In this study, using our recently developed nonlinear model, the pressure dependent (PD) attenuation and Cs of a bubbly medium are calculated and classified in terms of the resonance frequency (fr) of the MBs for (0.5fr<fr<4fr) and for pressures between 1 kPa<P<2 MPa. The propagation of the focused waves was simulated for each classified case using FEM simulation of the Helmholtz model. Results show that the presence of the MBs may shilled the target and eliminate the focal zone especially for frequencies close to fr. We show that, by choosing the optimal frequency (e.g., PD fr), the attenuation of the pre-focal region can be minimized while the attenuation at the focus maximized. This restores the focal pressure with minimum loss and decreases the pre and post-focal MB activity which may result in a better imaging or treatment efficacy.

5aBAb21. Study on ultrasound irradiating conditions for apoptosis induction in the rat glioma cell line C6. Sugamata Hiroaki, Satou Takaaki, and Takeuchi Shinichi (Dept. School of Eng., Toin Univ. of Yokohama, 1614 Kurogane-cho, Aoba-ku, Yokohama, Kanagawa 225-8503, Japan, b24c027o@ust.toin.ac.jp)

We're paying attention on ultrasound irradiation the cancer treatment method using apoptosis of a cell self-destruction program. We think that our method can be expected as a new effective and safety treatment method, if the cells were induced into apoptosis by using ultrasound. In this experiment, ultrasound irradiation experiments, to rat glioma cell line C6 were performed. The experiments were performed by using the standing wave type ultrasound irradiation system at frequency of 150 kHz. Also, we estimated the apoptosis induction after ultrasound irradiation by using Annexin V Assay Kit. We examined apoptosis induction of the cells after 8 hours of ultrasound irradiation. The fluorescence of FITC and PI was observed in the same cells, when ultrasound was irradiated to the cells by applying the continuous sinusoidal wave with voltage of 150 V_{pp}. It was suggested that the cells wave in the late stage of apoptosis and the stage of necrosis. When irradiated with applied voltage of 213 V_{pp}, only the fluorescence of FITC could be observed in the same cells. It was suggested an early stage of apoptosis. We're planning to perform the experiment, using the focused ultrasound irradiation system at frequency of 1.75 MHz in near future.

5aBAb22. Time delay spectrometry methods for broadband characterization of plastics and tissue-mimicking materials for high intensity therapeutic ultrasound. Subha Maruvada, Yunbo Liu (U.S. Food and Drug Administration, 10903 New Hampshire Ave., Bldg. WO 62-2222, Silver Spring, MD 20993, subha.maruvada@fda.hhs.gov), Paul Gammell (Gammell Appl. Technologies, Exmore, VA), and Keith Wear (U.S. Food and Drug Administration, Silver Spring, MD)

Because of the nonlinear nature of High Intensity Therapeutic Ultrasound (HITU), tissues and tissue-mimicking materials must be characterized over a broad frequency band. Time Delay Spectrometry (TDS) measures the frequency response of a system or sample by applying a swept frequency source and using a tracking receiver to select only those through-transmission signals that have the desired time delay. Due to the greater time-bandwidth product, TDS provides considerable signal-to-noise improvement over Fourier transform techniques. This system has been used previously to provide broadband measurements of complex hydrophone sensitivity. We

have further refined our TDS system to provide frequency-dependent attenuation and phase velocity characterization of materials over the range from 1 MHz to 19 MHz. We have compared the results to predicted dispersion curves based on the Kramers-Kronig relations. The measurements of low-density polyethylene (LDPE) and a HITU-compatible tissue-mimicking material agree to within 1% for both TDS and PR acquisition methods, but the TDS measurements provide superior bandwidth over the PR. Measured phase velocity agrees to within 1% with Kramers-Kronig-based dispersion curves from 1 MHz to 19 MHz.

5aBAb23. Targeted phase-change contrast agents for intracellular delivery to breast cancer tumors: An *In vitro* study in MDA-MB-231 breast cancer cells. Kyle P. Hadinger, Joseph P. Marshalek (Medical Imaging, Univ. of Arizona, Tucson, AZ), Paul S. Sheeran (Medical Biophys., Univ. of Toronto Stonybrook, Toronto, ON, Canada), Pier Ingram, Russell S. Witte (Medical Imaging, Univ. of Arizona, Tucson, AZ), Paul A. Dayton (Biomedical Eng., Univ. of North Carolina, Chapel Hill, NC), and Terry O. Matsunaga (Medical Imaging, Univ. of Arizona, PO Box 245067, 1609 N. Warren St., Tucson, AZ 85724, tmatsunaga@radiology.arizona.edu)

We have found that folate-targeted, ultrasound-activated phase-change contrast agents (PCCAs) composed of volatile perfluorocarbons are able to bind and internalize within breast cancer cells prior to vaporization into intracellular microbubbles. We here report studies assessing internalization, stability to thermal activation, and ultrasound activation of PCCAs formed by condensation of dodecafluoropentane (DDFP), decafluorobutane (DFB), and octafluoropropane (OFP). Using a Zonare "z.one-ultra" ultrasound system and L14-5sp probe operating in B-mode at 9 MHz (frame rate 13 Hz) to activate PCCAs and imaging with a Leica confocal microscope, we studied delivery in cultured MDA-MB-231 breast cancer cells with folate-targeted PCCAs. The number of internalized bubbles pre- and post- insonation were measured as a function of: 1) PCCA incubation times (i.e 10 mins, 60 mins); 2) ultrasound exposure (4 passes vs. 12 passes); and 3) mixtures of OFP, DFB, and DDFP. DFB:DDFP PCCAs, 60 minute incubations, and 12 passes offered a promising combination of stability and activation, with post-ultrasound to pre-ultrasound microbubble ratios approximating 3.4 (microbubbles pre-ultrasound per cell = 0.68 +/- 0.49, microbubbles post-ultrasound per cell = 2.35 +/- 0.79). Results indicate that intracellular delivery of targeted-PCCAs can potentially offer a new methodology for clinical application with ultrasound-mediated intracellular imaging of breast cancer.

5aBAb24. Micro ultrasound motor with coiled stator for rotationally driving of ultrasound beam of Intravascular ultrasound imaging system. Shinichi Takeuchi (Dept. of Biomedical Eng., Toin Univ. of Yokohama, 1614 Kurogane-cho, Aoba-ku, Yokohama 225-8503, Japan, shin1@toin.ac.jp), Keisuke Kurita, Chouyuu Uehara (Dept. of Biomedical Eng., Toin Univ. of Yokohama, Yokohama, Kanagawa, Japan), and Seiya Ozeki (Dept. of Medical Care Technol., Tsukuba Int. Univ., Tsuchiura, Ibaraki, Japan)

For the purpose of performing radial scanning by rotating ultrasound beam in intravascular ultrasound imaging system for diagnosing the state of thrombus and stenosis in a blood vessel such as the coronary arteries of the heart, we are developing micro Coiled Stator UltraSound Motor (CS-USM). Three types of coiled stators with materials of titanium, stainless steel or copper were prepared. These coiled stators were 5 wound acoustic waveguide with 0.5 mm width and 0.1 mm thickness. Size of coiled stator is 0.9 mm diameter and 3 mm length. PZT ceramic vibrator or hydrothermal PZT film vibrator was employed. When two vibrators-type inner rotor CS-USM having stainless steel coiled stator were driven by voltage of $30V_{pp}$ at 311 kHz, rotational speed was 3500 rpm, torque was 0.34μNm. Maximum rotational speed of one vibrator-type inner rotor CS-USM having hydrothermal PZT film was 1100 rpm. Maximum rotational speed of one vibrator-type inner rotor CS-USM with titanium or stainless steel coiled stator was 2400 rpm. Rotational speed of the one vibrator-type inner rotor CS-USM with copper coiled stator driven by same conditions was 3200rpm. Maximum rotational speed of two vibrator-type outer rotor CS-USM with stainless steel coiled stator and PZT ceramic vibrators was 1400 rpm.

5a FRI. AM

5aBAb25. Acoustic characterization of stethoscopes using auscultation sounds with synchronized electrocardiography recordings as test signals. Lukasz J. Nowak (Inst. of Fundamental Technolog. Res., Polish Acad. of Sci., ul. Pawinskiego 5B, Warszawa 02-106, Poland, lnowak@ippt.pan.pl) and Karolina M. Nowak (Dept. of Endocrinology, Ctr. of Postgraduate Medical Education, Warsaw, Poland)

The modern stethoscopes implement different technologies and features in order to better transmit the sounds from the inside of a body of a patient to the doctor's ears. However, the objective evaluation of the actual influence of various solutions on the acoustic parameters of the diagnostic device is difficult. This is due to the fact, that the acoustic coupling between the body of an auscultated patient and the chest piece of a stethoscope can significantly alter the determined parameters, and due to the variable and noisy nature of the auscultation signals. The present study introduces a detailed methodology for acoustic characterization of the stethoscopes using the auscultation sounds as test signals. Information obtained from additional, synchronized electrocardiography measurements is used to extract short, specific fragments of recordings, defined as acoustic events. Analysis of the spectral characteristics of many acoustic events allows to compare the acoustic properties of various stethoscopes and to estimate the measurement uncertainty. The determined statistical acoustic parameters of the auscultation signals are presented, showing relatively large variation even within a single patient examination. The example results concerning various commonly used acoustic stethoscopes are introduced, showing discrepancies between the declared and the actual acoustic properties.

5aBAb26. Discrepancies in the reproduction of the spectral characteristics of the acoustic stethoscopes by the electronic stethoscopes and potential diagnostic consequences. Karolina M. Nowak (Dept. of Endocrinology, Ctr. of Postgraduate Medical Education, Ceglowska 80 St., Warsaw 01-809, Poland, karolina.nowak@ippt.pan.pl) and Lukasz J. Nowak (Inst. of Fundamental Technolog. Res., Polish Acad. of Sci., Warsaw, Poland)

The electronic stethoscopes implement selectable low- and high-pass digital filters in order to mimic the declared acoustic characteristics of the bell and diaphragm chest pieces of the acoustic stethoscopes. However, the question arises if such a simple approach can exactly reproduce the actual, complex spectral characteristics of the acoustic stethoscopes, which are strongly affected by the acoustic coupling between the body of a patient and the chest piece? The present study introduces the results of the relevant experimental investigations. The acoustic characteristics of four different electronic stethoscopes operating in bell and diaphragm modes were determined during the heart auscultation examinations combined with the synchronized ECG recordings. The heart sounds were recorded using a microphone placed in one of the earpieces, while the ECG signals were used to extract interference-free fragments corresponding to single heartbeats. The averaged spectra of a large numbers of such fragments were used to determine the acoustic parameters of the investigated devices. The obtained results are compared to the corresponding results obtained for the acoustic stethoscopes, showing significant discrepancies. The potential diagnostic consequences of such discrepancies are discussed based on the available auscultation guidelines for physicians, which do not take into account the described effects.

5aBAb27. Study on measurement technique for acoustic cavitation using cavitation bubble signals. Takeyoshi Uchida, Masahiro Yoshioka, Youichi Matsuda, and Ryuzo Horiuchi (NMIJ, AIST, AIST Tsukuba Central3 1-1-1 Umezono, Tsukuba, Ibaraki 305-8563, Japan, takeyoshi.uchida@aist.go.jp)

Acoustic cavitation shows an excellent effect on ultrasonic cleaning and it also has prospects of application to cancer treatment. However, further study is essential to solve the problems such as low yield ratio and safety for human body because too much cavitation can damage the target. In this respect, quantitative evaluation of the generated cavitation is required. The mechanical index has been used so far but it does not reflect the quantity of the generated cavitation. Thus, we have been studying a new measurement technique by using the signal generated from cavitation bubbles, frequency spectrum of which is composed of broadband noise and subharmonics. We have reported that broadband noise has potential as an index for the amount of the generated cavitation. In this presentation, we experimentally

investigated the level change of both broadband noise and subharmonics by increasing the acoustic pressure. We found that these two components are observed at the different level of acoustic pressures, probably corresponding to the different kinetic state of the cavitation bubbles.

5aBAb28. Measurement of volumetric heat capacity of biological tissues heated by ultrasound exposure. Mariko Sugiyama, Hiroaki Kanayama, and Iwaki Akiyama (Medical Ultrasound Res. Ctr., Doshisha Univ., 1-3 Tatara-Miyakodani, Kyotanabe, Kyoto 610-0394, Japan, dmp1029@mail4.doshisha.ac.jp)

The authors have proposed a simple method of measuring volumetric heat capacity Cv from the small temperature rise of biological tissues heated by a short time ultrasound exposure. In this method, Cv is estimated by substituting the measured values of temperature rise speed and the acoustic parameters of specimen to the bio-heat transfer equation. The feasibility of the proposed method was studied by a block of porcine muscle as a specimen. In the experiments, it was heated by ultrasound exposure by the focusing concave transducer of diameter of 25 mm and focal length of 32.5 mm with a central hole of 2.0mm in diameter. The temperature rise was measured by a type K thermocouple which was inserted into the hole through the transducer. The measured value of temperature rise speed was $0.7°/s$, by an ultrasonic exposure of 2 MHz in frequency and 8.0 W/cm^2 in intensity. As a result, the value of Cv was measured as 3.6 $J/°C•cm^3$ which was approximately agreed with the value of Cv of porcine muscle referred as 3.7 $J/°C•cm^3$. [This study was supported by MEXT-Supported Program for the Strategic Research Foundation at Private Universities, 2013-2017.]

5aBAb29. Intracellular Ca^{2+} increase in endothelial cells induced by shock wave irradiation. Toru Takahashi (Graduate School of Sci. and Eng., National Defense Acad. of Japan, Hashirimizu 1-10-20, Yokosuka, Kanagawa 239-8686, Japan, em54026@nda.ac.jp), Keiichi Nakagawa (Graduate School of Eng., The Univ. of Tokyo, Bunkyo-ku, Japan), Tada Shigeru, and Akira Tsukamoto (Graduate School of Sci. and Eng., National Defense Acad. of Japan, Yokosuka, Japan)

Shock wave irradiations induce bioeffects including angiogenesis and bone healing. However, physical mechanisms underlying those bioeffects remain elusive. Extracorporeal shock wave treatment (ESWT) is one of well-known shock wave therapies. In ESWT, shock waves destroy kidney stones with cavitation and negative pressures. Energy flux density of shock waves in ESWT is around $0.1—1$ mJ/mm^2. In this study, shock waves with energy flux density of 0.001 mJ/mm^2, around 1/100 of ESWT, were irradiated on endothelial cells. In those cells, intracellular Ca^{2+} increase, i.e. a typical physiological response, was observed. Along the intracellular Ca^{2+} increase, extracellular fluid flow was not present. As well, both plasma membrane permeabilization and cell detachment were not detected. On the other hand, actin cytoskeletons, which link extracellular substrates and cellular components, were suggested to involve in the intracellular Ca^{2+} increase. Thus, it was supposed that shock waves initiated intracellular Ca^{2+} increase by loading forces on actin cytoskeletons. Although some previous studies have suggested that extracellular fluid flow could involve in shock wave induced bioeffects, intracellular Ca^{2+} increase seems to rely on extracellular fluid flow.

5aBAb30. Multimodal measurement of acoustic properties in cartilage tissue. Naotaka Nitta, Akio Kaya, Masaki Misawa, Koji Hyodo (National Inst. of Adv. Industrial Sci. and Technol. (AIST), 1-2-1 Namiki, Tsukuba 305-8564, Japan, n.nitta@aist.go.jp), and Tomokazu Numano (Tokyo Metropolitan Univ., Tokyo, Japan)

In order to detect the disease and degeneration of cartilage early, evaluations of acoustic properties are expected, but the physical meaning should be clarified. On the other hand, we have proposed a multimodal method using ultrasound (US) and magnetic resonance (MR) images for *in vivo* SOS measurement. In this study, the method was applied for the tissue-engineered cartilage, and the SOS was compared with T1, T2, apparent diffusion coefficient (ADC), attenuation coefficient (ATT), and Young's modulus (YM). All procedures were approved by the Animal Care and Use Committee of our institution. 12-week-old male rats were used, and five different

types of tissue-engineered cartilage was subcutaneously implanted in the back of rat (n = 30). Under anesthesia, MR and US images including same cross-sections were acquired separately, by using a MR imager of 2 Tesla and a US device of 13 MHz. At this time, T1, T2, ADC were also measured. The SOS was determined by the thickness measurement using the MR image and the time-of-flight measurement using the US image. After that, the ATT and YM were measured by using the extracted samples of tissue-engineered cartilage. As a result, the SOS exhibits negative correlations with T1, T2, and ADC (R^2 = 0.30, 0.93, 0.82), and positive correlations with ATT and YM (R^2 = 0.98, 0.87). These results suggest the SOS of tissue-engineered cartilage strongly reflects the elasticity. Consequently, the physical meaning of SOS was clarified through the multimodal measurement.

5aBAb31. A novel visualization method of bubble cavitation caused by infinitesimal amount of microbubbles. Ren Koda, Yoshiki Yamakoshi, Takumu Origasa, Toshitaka Nakajima (Grad. School of Sci. and Technol., Gunma Univ., 1-5-1 Tenjin-cho, Kiryu-shi, Gunma 376-8515, Japan, koda@gunma-u.ac.jp), and Takahito Nakajima (Grad. School of Medicine, Gunma Univ., Maebashi-shi, Japan)

Microbubbles modified with targeting ligands have a potential in early diagnosis by ultrasound because the targeted-bubble has both a selective property for accumulation in diseased area and a high degree of echogenicity. However, conventional ultrasonic imaging methods, such as B-mode and ultrasound color Doppler method, do not have enough sensitivity for microbubbles when only very small amount of microbubbles exists in ROI. In order to detect infinitesimal amount of microbubbles, we proposed a novel visualization of bubble cavitation signal (BCS) induced by high intensity US irradiation. In this method, high intensity US (h-US) is irradiated to microbubble with a fixed time delay after introducing an imaging US (i-US) and BCS is detected by power Doppler imaging unit. Due to the time delay of two waves, BCS appears on power Doppler image as a vertical line image (T-image) whose intensity is modulated by the amplitude of BCS. By this method, BCS measurement which is free both for direct propagated h-US and Doppler signal caused by bubble motion can be achieved. We detected the infinitesimal amount of bubbles (Sonazoid) with the concentration of 8×10^{-8} μL/ml, which is 100 times lower than minimum detectable sensitivity of B-mode observation.

5aBAb32. Focused ultrasound therapy of cervical lymph nodes in rats for alleviating multiple sclerosis symptoms. Anthony Podkowa, Rita J. Miller (Dept. of Elec. and Comput. Eng., Univ. of Illinois at Urbana-Champaign, 405 N Mathews, Urbana, IL 61801, tpodkow2@illinois.edu), Robert W. Motl (Kinesiology and Community Health, Univ. of Illinois at Urbana-Champaign, Urbana, IL), and Michael L. Oelze (Dept. of Elec. and Comput. Eng., Univ. of Illinois at Urbana-Champaign, Urbana, IL)

Multiple sclerosis (MS) is a prevalent neurological disease among adults worldwide. Disease modifying therapies are only modestly effective for slowing long-term progression of pathological and disability outcomes. One approach to treating MS is to reduce the burden of lymphocytes entering the central nervous system. We hypothesized that using focused ultrasound (FUS) to heat the cervical lymph nodes in rats with EAE, an animal model of MS, would alleviate symptoms of EAE by reducing lymphocyte burden. EAE was induced in rats through injection of myelin oligodendrocyte glycoprotein (MOG) and EAE disability scores were recorded over 21 days post injection. The cervical lymph nodes of rats with EAE were treated at day 9 and day 12 post MOG injection using FUS to elevate the temperature to 43-44 °C for 20 minutes. On average the EAE remittance score for FUS treated rats was 1.14 ± 0.48 versus 0.33 ± 0.27 for sham treated rats. These differences were statistically significant (p = 0.037). Remittance of the EAE disability scores were highly correlated with the last therapy application. Therefore, FUS treatment of cervical lymph nodes in rats with EAE resulted in a significant reduction in EAE score, which was not observed in sham treated rats.

5aBAb33. Shear wave elastography reveals dispersion regimes in porous materials. Johannes Aichele, Chadi Zemzemi, Maxime Lescanne, Stefan Catheline (Labtau, INSERM, Cours Albert Thomas 152, Lyon 69003, France, johannes.aichele@inserm.fr), and Philippe Roux (ISTERRE, GIÈRES, France)

Ultrasound shear wave elastography is a well established tool for characterization of biologic tissues. While it has found application in various medical disciplines such as oncology and urology its feasibility for pneumology still has to be shown. We provide experimental results of ultrasound shear wave elastography of porous materials in phantoms and ex-vivo lung tissue. Phantom foams immersed in water show a strong phase velocity dispersion with increasing frequency. Two regimes can be identified in the dispersion curves in the investigated frequency range from 50 to 700 Hz. These experimental results are in reasonable agreement with the theory of porous materials. The results from water-filled phantom foams are compared to gelatin-filled phantom foams. Finally, the phantom study is compared to surface waves dispersion curves of porcine lungs obtained with an ultra-fast optical camera and the first results of ex-vivo shear wave elastography in porcine lung. Ultrasound shear wave elastography is standard in non-porous organs such as muscle, liver, and breast tissue. Its application to porous materials could offer a noninvasive and nonionizing alternative for lung characterization.

5aBAb34. Reconstruction of sound speed distributions from pulse-echo data. Anthony Podkowa, Mert Hidayetoglu, Chunxia Yang, Michael Oelze, and Weng C. Chew (Elec. and Comput. Eng., Univ. of Illinois at Urbana-Champaign, 1009 W. Clark St. Apt. 205, Urbana, IL 61801, tpodkow2@illinois.edu)

Ultrasonic imaging is perhaps the most ubiquitous form of biomedical imaging currently employed in the clinical setting due its safety, affordability, and real-time imaging capabilities. However, the poor image quality of standard brightness mode (B-Mode) images has inspired research into alternative acoustic imaging modalities, such as sound speed imaging. Historically, sound speed tomography has been employed in transmission-mode because of the limited k-space coverage available in backscattered data. However, such imaging techniques require specialized hardware configurations, and are often only limited to tissue such as the breast where full angular coverage is available. A recent approach [Jaeger *et al*., UMB, 41: 235-250, 2015] has demonstrated the feasibility of sound speed imaging using only pulse-echo data, which only requires a single array probe, commonly available in a clinical setting. This approach allows for complementary k-space coverage by utilizing relative temporal shifts between successive steered plane wave transmissions. Motivated by these successes, we present simulated and experimental pulse-echo sound speed reconstructions using three different methods: a modified Jaeger-like approach, the Born iterative method (BIM), and the distorted Born iterative method (DBIM).

5aBAb35. Experimental and theoretical studies on acoustic droplet vaporization. Krishna N. Kumar, Mitra Aliabouzar, and Kausik Sarkar (George Washington Univ., 801 22nd St. NW, Washington, DC 20052, sarkar@gwu.edu)

Phase shift nanodroplets are better alternatives to microbubbles due to their enhanced stability and smaller size distribution. These nanodroplets undergo phase transition from liquid to highly echogenic gaseous state under acoustic excitation through a process termed acoustic droplet vaporization (ADV). In this study, we synthesized lipid-coated perfluoropentane (PFP)-filled nanodroplets via sonication and mechanical agitation methods. We investigated the ADV threshold of these nanodroplets as a function of acoustic parameters such as excitation frequency and pulse repetition period (PRP). Experiments were performed at frequencies 2.25, 5, and 10 MHz. The acoustic signature of droplet vaporization was observed to be a broadband signal at all the studied frequencies. The ADV threshold was studied by increasing the excitation pressure in steps of 200 kPa. The ADV threshold at 2.25 and 5 MHz was found to be 1.8 and 2.2 MPa, respectively. The scattered response from droplets were studied at different PRPs of 500 çs, 1 ms, and 10 ms; the results indicate that the scattered response is PRP-independent. Classical nucleation theory (CNT) was employed to determine the dependence of ADV threshold on parameters such as surface tension of the

5a FRI. AM

liquid, droplet size, degree of superheat of the droplet liquid, and excitation frequency.

5aBAb36. High frequency and high intensity ultrasonic transducers for medical applications using oriented piezoelectric thick films by hydrothermal method. Mutsuo Ishikawa (Medical Eng., Toin Univ. of Yokohama, 1614, Kurogane, Aobaku, Yokohama 2258503, Japan, m.ishikawa@toin.ac.jp), Yosuke Uchida (Medical Eng., Toin Univ. of Yokohama, Yokohama, Kanagawa, Japan), Marie Tabaru, Minoru Kurosawa (Tokyo Inst. of Technol., Yokohama, Kanagawa, Japan), Nobuyuki Kosuge, and Hideto Sugiyama (Medical Eng., Toin Univ. of Yokohama, Yokohama, Kanagawa, Japan)

Recently, high-frequency ultrasonic transducers are proposed for medical applications using piezoelectric crystal. In this case, it is known that the high piezoelectric constant and the higher limitation of vibration velocity are important for high power application. Normally, the piezoelectric ceramics has limitation due to maximum vibration velocity. Thus it is very difficult to radiate for high intensity ultrasound with nonlinear acoustics due to the limitation of maximum vibration velocity. Therefore, we tried to develop a very high intensity and high frequency ultrasonic transducer using hydrothermal $KNbO_3$ thick film. The piezoelectric constant was over 80 pm/V and the vibration velocity of thickness mode of $KNbO_3$ thick film was 2.5 m/s or over at 20 MHz. This maximum vibration velocity is twice of reported piezoelectric materials. Then, a radiation sound-pressure from the prototype ultrasonic transducer with $KNbO_3$ thick film was measured. The result of radiation ultrasound was 4 MPa or over in a water. It is very high intensity at high frequency. Consequently the hydrothermal $KNbO_3$ thick film is usefulness for medical applications. The reason for why it is able to radiate for high intensity ultrasound at high frequency will be discussed in this presentation.

5aBAb37. Sensitivity study of high-frequency ultrasound to skin cancer using micro-tumor mimicking phantoms. Nicole Cowan (BioTechnol., Utah Valley Univ., 1294 Parkside Dr., Saratoga Springs, UT 84045, ncowan18@gmail.com), Zachary A. Coffman (Biology, Utah Valley Univ., West Jordan, UT), Robyn K. Omer (Botany, Utah Valley Univ., Orem, UT), Caitlin Carter (BioTechnol., Utah Valley Univ., Orem, UT), and Timothy E. Doyle (Phys., Utah Valley Univ., Orem, UT)

Skin cancer is the most common cancer in the United States. An estimated one in five Americans will be diagnosed with skin cancer in their lifetime with approximately 145,000 new cases each year. Mohs surgery (Mohs micrographic surgery) is a surgical technique used to treat skin cancer. During Mohs surgery, thin layers of skin tissue are excised from the cancer site and examined until only cancer-free tissue remains. The examination of excised layers is tedious and time-consuming (4-5 hours). A rapid assessment tool for detecting cancerous tissue in excised margins would prove beneficial for both dermatologists and their patients. High-frequency (HF) ultrasound (20-80 MHz) has been explored as a rapid assessment tool for such cases. The purpose of this study was to test the sensitivity of HF ultrasound to micro-tumors present in skin cancer margins using histology mimicking phantoms. Phantoms were created from distilled water, agarose powder, 10X TBE stock solution, and polyethylene microspheres to simulate micro-tumors. Pitch-catch and pulse-echo measurements were acquired using a high-frequency ultrasound system with glycerol as the coupling agent. The phantom results show consistent increases in attenuation when testing positions containing micro-tumors, confirming the sensitivity of HF ultrasound to skin cancer pathologies.

FRIDAY MORNING, 2 DECEMBER 2016　　　　　　　SOUTH PACIFIC 4, 7:45 A.M. TO 12:00 NOON

Session 5aEA

Engineering Acoustics: Applied Acoustics I

Bernhard R. Tittmann, Cochair
Eng. Sci. & Mechanics, Penn State University, 212 EES Bldg., University Park, PA 16802

Masatonakayama, Cochair
Ritsumeikan Univ., Kyoto, Japan

Contributed Papers

7:45

5aEA1. Ultrasound for nuclear reactors. Bernhard R. Tittmann, Brian Reinhardt, and Andrew Suprock (Eng. Sci. & Mech., Penn State Univ., 212 EES Bldg., University Park, PA 16802, brt4@psu.edu)

Ultrasonic methods offer the potential for Structural Health Monitoring of critical components in nuclear reactors. These efforts have been limited by ultrasonic transducers incapable of performance under high temperatures and/or irradiation conditions. Here we report on piezoelectric transducers designed, fabricated, tested and optimized to perform in harsh environments. Test capsules with piezoelectric transducers were fabricated with Aluminum Nitride (AlN), Zinc Oxide (ZnO), and Bismuth Titanate (BiTi) as the active elements. Measurements were performed in the MIT Reactor for 18 months. The transducers experienced an integrated neutron fluence of approximately 8.68 E+20 n/cm^2 for n >1 MeV, temperatures in excess of 420 °C, and a gamma fluence of 7.23 Gy/cm^2. The AlN transducer acoustically coupled to a Kovar cylinder gave acceptable pulse-echo data throughout the test. We show a summary of the test results. Thus the feasibility of ultrasonic transducers in a nuclear reactor has been established and opens the door to leave-in-place sensors for in-reactor conditions and materials. [The authors gratefully acknowledge support from the Department of Energy under the ATR-NSUF program.]

8:00

5aEA2. Writer recognition with a sound in hand-writing. Daichi Asakura, Takanori Nishino, and Hiroshi Naruse (Graduate School of Information Eng., Mie Univ., 1577 Kurimamachiya-cho, Tsu, Mie 514-8507, Japan, asakura@pa.info.mie-u.ac.jp)

Hand writing sound is one of noises existing in the real environment. Generally speaking, noises are meaningless; however, some noises, e.g., footsteps and a road noise, are effective to recognize circumstances. Hand

writing sound also has some information to recognize a writing character. Previous study addressed the recognition of handwritten numeric characters by using writing sound, and the average recognition rate was 88.4%. We advanced previous research and examined the possibility of writer recognition from writing sound. We asked our subjects to write three target names. When subjects wrote on paper by a ballpoint pen, writing sounds were recorded. Three subjects chosen as the target name also wrote their own names under the different conditions, e.g., writing slowly and writing in another day. In our experiments, writer recognition was performed by the hidden Markov model. Acoustic features were mel-frequency cepstrum coefficients (MFCC), delta coefficients, and delta-delta (acceleration) coefficients. Experimental results showed that the writer recognition is possible; however, problems still exist such as how to deal with the difference between the writing conditions. In future works, it is necessary to increase learning data and to improve the recognition method.

8:15

5aEA3. Effect of the horizontal panning on sense of presence in three-dimensional audio system based on multiple vertical panning. Toshiyuki Kimura (Faculty of Eng., Tohoku Gakuin Univ., 1-13-1, Chuo, Tagajo, Miyagi 985-8537, Japan, t-kimura@m.ieice.org) and Hiroshi Ando (Ctr. for Information and Neural Networks, National Inst. of Information and Communications Technol., Suita, Osaka, Japan)

We previously proposed a three-dimensional (3D) audio system using the multiple vertical panning (MVP) method to develop a 3D audio system that matches a multi-view 3D video display system (REI display). In this paper, in order to apply our proposed method to the teleconference system, we performed two audio-visual psychological experiments and evaluated the effect of the horizontal panning on the sense of presence in our proposed method. We found that the performance is maintained when four loudspeakers were placed at the position except the edges of the display.

8:30

5aEA4. Design of a bullet beam pattern of an ultrasound transducer with shaded electrodes and a multifocal lens. Yongrae Roh, Seongwon Jang, Euna Choi, and Yeonsue Park (School of Mech. Eng., Kyungpook National Univ., 80 Daehakro, Bukgu, Daegu 41566, South Korea, yryong@knu.ac.kr)

Ultrasonic imaging transducer is required to compose a beam pattern of a low sidelobe level and a small beam width to achieve good image resolution. Normal ultrasound transducers have many channels along its azimuth, which allows easy formation of the sound beam. However, they have no control of the beam pattern along their elevation. In this work, a new method is proposed to manipulate the beam pattern by using an acoustic multifocal lens and shaded electrodes in the transducer. The shading technique splits an initially uniform electrode into several segments. For a given elevation width and frequency, the optimal pattern of the split electrodes was determined to achieve the lowest sidelobe level. The requirement to achieve a small beam width with a long focal region was satisfied by employing an acoustic lens of three multiple focuses. The geometry of the multifocal lens was also optimized. For the optimization, a new index was devised to evaluate the on-axis response: focal region ratio = focal region/minimum beam width. The larger was the focal region ratio, the better was the beam pattern. The validity of the design has been verified through fabricating and characterizing an experimental prototype of the transducer.

8:45

5aEA5. Fast measurement of the temperature distribution in a blast furnace using impulse response measurement. Gottfried Behler, Jan-Gerrit Richter (Inst. of Tech. Acoust., RWTH Aachen Univ., Kopernikusstraße 5, Aachen D-52074, Germany, gkb@akustik.rwth-aachen.de), and Sebastian Buzga (Z&J, IMI Critical Eng., Düren, Germany)

The temperature profile on top of a blast furnace gives meaningful information about the heat distribution and the activity of the ballast below. Since a direct measurement using temperature probes is both complicated due to the filling procedures and expensive due to the repeated damage of the sensors. An indirect measurement of the temperature is given by the measurement of the propagation delay for sound which by a known distance

reveals the average temperature of the gas in between source and receiver position. The temperature distribution is then derived by a tomographic approach in one plain above the blast furnace with different source and receiver positions. Basically, this method is not new and it is used with stochastic (pneumatic excitation) signals in many applications. The typical measurement time for a single temperature profile takes about 20-30 seconds. It is obvious that the temperature distribution will not be stable for such a long time; hence, the measurement results are not very much reliable. The background noise in the blast furnace and the excitation signal are of the same type (noise), hence the correlation of the measured signal very often fails. In this paper, a new electroacoustic method using interleaved swept sine excitation will be described.

9:00

5aEA6. Defect position and size estimation in billet by ultrasonic transmission method using a small number of transducers. Ryusuke Miyamoto (Graduate School of Systems and Information Eng., Univ. of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan, miyamoto@aclab.esys.tsukuba.ac.jp), Koichi Mizutani, Tadashi Ebihara, and Naoto Wakatsuki (Faculty of Eng., Information and Systems, Univ. of Tsukuba, Tsukuba, Japan)

For a steel billets inspection, we have proposed defect detection and size estimation from time-of-flight (TOF) profile by ultrasonic transmission method with linear scanning. In this method, a defect is detected by deviation of TOF of transmitted wave caused by diffraction at the defect. In addition to defect detection, defect size can be estimated from deviation of TOF because TOF is affected by defect size. However, deviation of TOF is also affected by the depth of the defect, which cannot be precisely estimated only by linear scanning of a pair of transducers. In this study, we propose defect position and size estimation using multiple pairs of transducers and validity of the method was evaluated. By using multiple pairs of transducers, TOF of inclined paths, where transducers is not facing each other are also obtained as well as parallel paths, so that the position of defect can be estimated. As a result, the defect position could be estimated by using two pairs of transducers and defect size could also be estimated correctly compared with conventional linear scanning method.

9:15

5aEA7. Shape optimization of reactive mufflers using threshold acceptance and finite element methods. Abdelkader Khamchane (Laboratoire de Mécanique Matériaux et Energétique (L2ME), Université Abderahman Mira de Bejaia, Rte. de Targa Ouzemour, Béjaia 06000, Algeria, abdelkader.khamchane@yahoo.fr), Youcef Khelfaoui, and Brahim Hamtache (Laboratoire de Mécanique Matériaux et Energétique (L2ME), Université Abderahman Mira de Bejaia, Bejaia, Bejaia, Algeria)

Recently, research on the acoustic performance of reactive mufflers under space constraint becomes important. The attenuation performance of single and double expansion-chambers under space constraint is presented in this paper. A shape optimization analysis is performed using a novel scheme called Threshold Acceptance (TA), the best design obtained by the shape optimization method are analysed by Finite Element Method (FEM). The acoustical ability of the mufflers obtained is than assessed by comparing the FEM solution with the analytical method. Results show that the maximal STL is precisely located at the desired targeted tone. In addition, the acoustical performance of mufflers with double expansion-chamber is found to be superiors to the other one. Consequently, this approach provides a quick and novel scheme for the shape optimization of reactive mufflers.

9:30

5aEA8. A mixed-order ambisonic scheme to improve performance for sound sources on the horizontal plane. Jiho Chang (Korea Res. Inst. of Standards and Sci., 267 Gajeong-ro, Yuseong-gu, Daejeon 34113, South Korea, chang.jiho@gmail.com)

This paper proposes a mixed-order ambisonic scheme called #P#L that improves performance in terms of estimation of the coefficient of spherical harmonics expansion, reconstruction of sound fields, and spatial resolution for sound sources located on the horizontal plane. Mixed-order ambisonics (MOA) selects appropriate components of spherical harmonics expansion instead of truncating up to a certain order as in higher-order ambisonics

5a FRI. AM

(HOA). Several schemes of MOA including the proposed scheme and HOA are compared each other in several ways. The results show that the proposed scheme has several advantages compared with the other MOA schemes. This scheme can be useful to further improve performance of microphone arrays if the number of the microphones is greater than what is required for HOA schemes or a non-uniform layout of microphones is used.

9:45–10:00 Break

10:00

5aEA9. Force feedback microelectromechanical microphones for high performance applications. Owain L. Boorman (ISVR, Univ. of Southampton, Bldg. 13, Highfield Campus, University Rd., Southampton, Hampshire SO17 1BJ, United Kingdom, olb1g11@soton.ac.uk), Nicholas Harris (ECS, Univ. of Southampton, Southampton, United Kingdom), and Matthew C. Wright (ISVR, Univ. of Southampton, Southampton, United Kingdom)

Microelectromechanical System (MEMS) condenser microphones are widely used because of their low cost, small size, high sensitivity, and wide bandwidth. For certain specialist applications, however, they are still outperformed by the best conventional condenser microphones, which have greater bandwidth and dynamic range, but at higher cost and larger size. The sensitivity, and hence, signal-to-noise ratio of smaller MEMS microphones can be increased by using two perforated back-plates instead of one. The maximum amplitude is limited by membrane excursion, which leads to nonlinearity and, ultimately, failure. The use of force feedback holds the promise of avoiding these problems by holding the membrane at its equilibrium position while measuring the force required to do so. Previous attempts to accomplish this using a Sigma-Delta modulator have had only limited success in terms of signal-to-noise ratio, bandwidth and stability. Instead we propose to use an Electro-Mechanical Phase Locked Loop (EMPLL) to overcome these limitations. We will present lumped-parameter and Finite Element models of the performance of such a microphone, and discuss the challenges associated with its fabrication. [This work was supported by Roke Manor Research Limited.]

10:15

5aEA10. Personalization of head-related transfer functions in the median plane based on spectral correction with pinna angle. Sakiko Mishima, Hajime Komatsu, Takahiro Fukumori, Masato Nakayama, and Takanobu Nishiura (Ritsumeikan Univ., 1-1-1, Noji-higashi, Kusatsu, Shiga 525-8577, Japan, is0188if@ed.ritsumei.ac.jp)

Head-related transfer functions (HRTFs) are studied to realize high-precision sound field reproduction systems. These systems usually utilize the HRTFs of other listeners because self-measurement of one's own HRTFs is a heavy burden. The localized sound accuracy of the listener, especially in the median plane, is degraded when the systems utilize the HRTFs of other listeners, which have different characteristics from the listener's own ones. Thus, simple personalization of HRTFs is necessary for solving this problem. In this research, we focus on personalization of HRTFs in the median plane on the basis of spectral correction with pinna angles as anthropometric parameters. We first made analyses of the relationship between pinna angles and the spectral envelopes of HRTFs in the median plane through an experiment with a dummy head with a replicated pinna made from silicone. From these analyses, the spectral envelopes were confirmed to be affected by pinna angles. We then proposed a method for personalization of HRTFs on the basis of spectral correction with the pinna angles. Objective and subjective evaluation experiments were carried out in order to demonstrate the effectiveness of the proposed method. As a result, we confirmed improvement in the performance of sound image localization.

10:30

5aEA11. Design of distributed mode loudspeaker through evolutionary structural optimization. Goki Shirouzu (Graduate School of Design, Kyushu Univ., 4-9-1 Shiobaru Minami-ku, Fukuoka, Fukuoka 815-8540, Japan, shirouzu.g.111@s.kyushu-u.ac.jp) and Toshiya Samejima (Faculty of Design, Kyushu Univ., Fukuoka, Fukuoka, Japan)

The distributed mode loudspeaker (DML) is a flat panel loudspeaker that consists of a thin plate as a diaphragm and several driving exciters. In designing the DML, a number of flexural resonance modes are encouraged to be excited. Thereby, complex vibration is produced in the diaphragm, and the sound field around the DML is diffuse at high frequencies. In the DML design procedures, choosing proper aspect ratio of the diaphragm, driving points, and suspension points are important to deliver flexural resonance modes as many as possible. In this research, the authors focus on the shape of diaphragm. The authors try to optimize the diaphragm shape that delivers more equalized interval between the natural frequencies of the diaphragm so that it does not have many degenerated modes. To achieve this, Evolutionary Structural Optimization (ESO) method, which is one of structural topology optimization methods, is adopted. The scheme is based on the idea that, by gradually removing inefficient materials, the structure shape evolves toward optimum one. This method enables us to select the best shape of the diaphragm from shapes given in the process. Trial designs demonstrate that the ESO method is effective in designing the diaphragm of the DML.

10:45

5aEA12. Spatial post-filter estimation for speech enhancement in the specific area using a pair of microphone arrays. Takuto Yoshimizu (Graduate School of Sci. and Technol., Ryukoku Univ., 1-5, Seta Oe-cho Yokotani, Otsu-shi, Shiga-ken 520-2194, Japan, t15m083@mail.ryukoku.ac.jp) and Akitoshi Kataoka (Faculty of Sci. and Technol., Ryukoku Univ., Otsu-shi, Shiga-ken, Japan)

Suppression methods using the beamformer and the Wiener post-filter by microphone array has been proposed. However, these methods are hard to suppress noise on the environment, such as the desired source is surrounded by noise. We propose the suppression method using a pair of microphone arrays arranged in different positions on the environment, such as the desired source is surrounded by noise. We can get different information of sound by using two microphone arrays. We make the post-filter using its different information of sound. This method enhance the desired source to be the sound source in the specific area. Therefore, we divide the room into nine areas by using two microphone arrays. By estimating the cross spectrum between two microphone arrays of each partition area, we make the post-filter to noise suppression. By the experiment results under actual environment, we confirmed that this method was higher in noise suppression performance than MVDR using one microphone array that is general beamformer.

11:00

5aEA13. Noise impact assessment on residential area near highway and noise control measures development. Ilya E. Tsukernikov, Igor L. Shubin, Leonid A. Tikhomirov, Natalia E. Schurova (Res. Inst. of Bldg. Phys., Odoevskogo proezd, h.7, korp.2, fl. 179, 21 Lokomotovny pr., Moscow 117574, Russian Federation, 3342488@mail.ru), and Ilya O. Tsukernikov (Moscow State Univ. of Printing Arts, Moscow, Russian Federation)

Currently, from 30 to 50% of urban population of the Russian Federation it is affected by the increased traffic noise. A similar situation is observed not only in big cities, but also in rather small settlements. Instrumental and calculated assessment of existing noise levels in the residential area "Zarechie" adjacent to a site of the Moscow Ring Road and having a soundproofing barrier on the border of the territory is performed in the paper. Acoustic situation on the territory without existing soundproofing barrier is modeled with use of the program "ARM Acoustics," developed by Russian Company "TECHNOPROEKT" and the effectiveness of the noise reducing by the screen is evaluated. A calculation for feasibility of a further increase in the height of the existing barrier is carried out. Noise maps are built which allowed to identify the main sources of acoustic discomfort in the residential area. The analysis of the impact on the acoustic environment of the building being built is conducted. Against this background options for protecting the area from noise are proposed.

11:15

5aEA14. Wideband audio systems for hearing aid devices. Michael Salameh (Starkey Hearing Technologies, 6600 Washington Ave. S., Eden Prairie, MN 55344, michael_salameh@starkey.com)

Multi-loudspeaker design has been one of the most widely used methods for improving the sound quality of audio systems. Some of the

electroacoustic design concepts defined for typical earphones can be used for other wideband miniature audio systems as well. However, wideband audio for hearing aid devices and smart earphones introduce new design factors. Efficiency and transducer size are among the factors that become important in the electroacoustic design. In this paper, the electroacoustic design factors for woofer-tweeter audio systems are revisited and discussed for hearing aid devices. Simulation and measurement results are presented and analyzed.

11:30

5aEA15. A broadband acoustic absorber based on phase-modulating reflective metasurfaces. Wenqi Wang, Junfei Li, Bogdan-Ioan Popa, and Steven Cummer (Elec. and Comput. Eng., Duke Univ., 3417 FCIEMAS, 101 Sci. Dr, Durham, NC 27708, wenqi.wang@duke.edu)

Most high performance acoustic absorbers exhibit outstanding absorption capabilities at resonant frequencies, thus their bandwidth are significantly limited. Acoustic metasurfaces enabled unprecedented wave manipulation with their planar profiles, subwavelength thicknesses, and large degree of design freedom. However, studies have been primarily focused on extraordinary wavefront shaping functionalities, and little explorations has been made on manipulating transmitted/reflected power of diffracted beams. By carefully tailoring the design and arrangement of acoustic metamaterial cells, enhanced absorption can be achieved for a broad range of frequencies. Building on our previous works on transmissive and reflective metasurfaces with exotic properties, including anomalous refraction, surface mode conversion, and extraordinary beam-steering (Xie *et al.* Nat. Commun. 2014), we present in this work a broadband acoustic absorber built with an expanded library of labyrinthine acoustic metamaterials. We demonstrate that by engineering the phase modulation profile of phase-modulating metasurfaces, incident waves can be effectively converted to surface modes that propagate along the air-metasurface interface, thus minimal reflection is observed in the far field. Our work extends functionalities of acoustic metasurfaces, demonstrates an alternative route to the design of broadband acoustic absorbers, and can be potentially applicable to noise control.

11:45

5aEA16. Intracochlear sound sensor-electrode system for fully implantable cochlear implant. Chuming Zhao, Katherine E. Knisely (Mech. Eng., Univ. of Michigan, 2350 Hayward, Ann Arbor, MI 48109, chumingz@umich.edu), Deborah J. Colesa, Bryan E. Pfingst, Yehoash Raphael (Kresge Hearing Res. Inst., Dept. of Otolaryngology-Head and Neck Surgery, Univ. of Michigan, Ann Arbor, MI), and Karl Grosh (Mech. Eng., Univ. of Michigan, Ann Arbor, MI)

We designed, fabricated, and tested an intracochlear sound sensor-electrode system consisting of an intracochlear sound sensor (ISS) and a 50μm Pt-Ir wire electrode. This system was designed to sense acoustic signals and transmit electrical stimuli inside the cochlea. Potential applications include acting as the front end of a fully implantable cochlear implant to treat sensorineural deafness or as a transducer in cochlear mechanics experiments. The ISS was comprised of an array of piezoelectric cantilevers and was built using micro-electrical-mechanical-system (MEMS) techniques. The ISS was tested in air and underwater to compare its functionalities to the numerical predictions. Later, the ISS was implanted in an anesthetized guinea pig and tested *in vivo*. A 90-100dB SPL pure tone acoustic excitation over 1-30kHz was delivered into the ear canal of the guinea pig, and 0.1-30μV was measured by the ISS. The sensed signal was linear, repeatable, and immune to the electrical interference from the extracochlear and intracochlear environment. When using the electrode, electrical auditory brainstem response (eABR) measurements were performed by sending a 25μs single pulse stimulus to the electrode. Stimulus amplitudes ranging in 220-400μA were found to evoke an eABR. These results show that the concept of sensing acoustic signals and transmitting electrical stimulation inside the living cochlea using one device is feasible.

Session 5aID

Interdisciplinary: Topical Meeting on Data Science and Acoustics I

Matthew G. Blevins, Cochair
U.S. Army Engineer Research and Development Center, 2902 Newmark Drive, Champaign, IL 61822

Andrew Christian, Cochair
National Institute of Aerospace, 100 Exploration Way, Hampton, VA 23666

Joonhee Lee, Cochair
Durham School of Architectural Engineering, University of Nebraska - Lincoln, 1110 S. 67th Street, Omaha, NE 68182-0816

Hiroshi Sato, Cochair
Dept. of Information Technology and Human Factors, Natl. Inst. of Advanced Industrial Sci. and Tech., Tsukuba, Japan

Chair's Introduction—8:35

Invited Papers

8:40

5aID1. High performance and robust audio search. Avery Wang (Shazam, 2114 Broadway St., Redwood City, CA 94063, avery.wang@shazam.com)

In this talk, I will give an overview of the Shazam audio recognition technology. The Shazam service takes a query comprised of a short sample of ambient audio (as little as 2 seconds) from a microphone and searches a massive database of recordings comprising up to 40 million songs. The query audio may be distorted with significant additive noise ($<$0 dB SNR), environmental acoustics, as well as nonlinear distortions. The computational scaling is such that a query may cost as little as a millisecond of processing time. Previous algorithms could index hundreds of items, required seconds of processing time, and were less tolerant to noise and distortion by 20-30 dB SNR. In aggregate, the Shazam algorithm represented a leap of more than $1E+9$ in efficiency. I will discuss the various innovations leading to this result.

9:25

5aID2. Data visualization methods. Claudio Silva (New York Univ., 2 MetroTech Ctr., Rm 10.093, New York, NY 11201, csilva@nyu.edu)

Future advances in science, engineering, and medicine depend on the ability to comprehend the vast amounts of data being produced and acquired. Data visualization is a key enabling technology in this endeavor: it helps people explore and explain data through software systems that provide a static or interactive visual representation. Despite the promise that visualization can serve as an effective enabler of advances in other disciplines, the application of visualization technology is non-trivial. The design of effective visualizations is a complex process that requires understanding of existing techniques and how they relate to human cognition. For a visualization to be insightful, it needs to be both effective and efficient. This requires a combination of design and science to reveal information that is otherwise obscured. In this talk, we will give a general background for data visualization, and present some work on interactive visualization techniques and tools for a variety of analysis purposes, including signal processing, some of which can be repurposed for the analysis of audio data.

10:10–10:30 Break

10:30

5aID3. Primitives in time series mining: Algorithms and applications. Abdullah A. Mueen (Comput. Sci., Univ. of New Mexico, 1 University of New Mexico, Albuquerque, NM 87131, mueen@unm.edu)

Time series patterns are waveforms with properties useful for various data mining tasks such as summarization, classification and anomaly detection. In this talk, I present three types of time series patterns: Motifs, Shapelets, and Discords. Motifs are repeating patterns that repeat in seemingly random time series data; Shapelets are small segments of long time series characterizing their sources; Discords are anomalous waveforms in long time series that do not repeat anywhere else. I briefly discuss efficient algorithms to discover these patterns and present cases in mining data from robots, humans and social media. Cases include activity classification using accelerometer data, correlated clusters in social media data, and anomaly in astronomical data.

5aID4. The promises and realities of machine learning. Dan Valente (Chartbeat, 826 Broadway, 6th Fl., New York, NY 10003, dan@chartbeat.com)

These days, anyone with a laptop, Python, and a free afternoon can build an intelligent machine. In fact, if you have a bunch of data, go ahead and throw it into a deep neural net: your worries will be over! Machine learning will save us all! Unfortunately, those of us who spend our days fighting with these systems know that this is far from the truth. As anything created by humans, these algorithms have limitations, and there is no algorithmic substitute for careful thought. The gory details of techniques like Support Vector Machines, Neural Networks, and K-means clustering can be found in innumerable blog posts, texts, and online courses; we won't discuss the details here. Instead, this talk will focus on the conceptual and mathematical framework under which all machine learning algorithms fall. We'll briefly discuss the workings of a few of the most popular modern techniques (#DeepLearning, anyone?) and how they fit into this framework. More importantly, we'll discuss some often-overlooked considerations that go into the design of these systems: human biases, transparency, and maintainability. Ultimately, we hope to provide a solid understanding of what these algorithms can do, what they can't do, and when they should or should not be used.

FRIDAY MORNING, 2 DECEMBER 2016　　　　　　　　　　HIBISCUS, 8:30 A.M. TO 11:30 A.M.

Session 5aMU

Musical Acoustics: General Topics in Musical Acoustics I

James P. Cottingham, Chair
Physics, Coe College, 1220 First Avenue NE, Cedar Rapids, IA 52402

Contributed Papers

8:30

5aMU1. Analysis of the relationship between muscle activity and acoustic features during trumpet play and the construction of a myoelectric visual feedback system. Megumi Satou (Library, Information and Media Studies, Univ. of Tsukuba, 1-2, Kasuga, Tsukuba, Ibaraki 305-8550, Japan, megumi@slis.tsukuba.ac.jp), Tetsuro Kitahara (Information Sci., Nihon Univ., Tokyo, Japan), Hiroko Terasawa, and Masaki Matsubara (Library, Information and Media Studies, Univ. of Tsukuba, Tsukuba, Ibaraki, Japan)

Breath control and lip vibration are crucial for a stable performance while playing the trumpet. We analyzed the differences in abdominal and orofacial muscle activity in acoustic features such as pitch (B♭2, F3, B♭3, F4, B♭4), intensity (pp, mf, ff), and duration (0.75, 6 s) during the preparation and sustain periods using surface electromyography in 11 amateur trumpeters. When the pitch was high, the activity of both muscles increased in both the preparation and sustain periods. However, when the intensity was high, the activity of both muscles increased only during the sustain period. Orofacial muscle activity was lower after tone production and abdominal muscle activity was higher after tone production than before tone production. In addition, we developed a visual feedback system that displays the muscle activities and acoustic features related to the produced sound as biofeedback can make learning performing technique efficient (LeVine, 1984). This system enables the player to objectively recognize the acoustic features and muscle activities related to sound. We aim to enhance performing technique by determining the relationship between the use of the body during a performance and the produced sound.

8:45

5aMU2. Sound has size: Stages of concert halls. James B. Lee (None, 6016 S. E. Mitchell, Portland, OR 97206, cadwal@macforcego.com)

Sabine characterized whole concert halls by ratio of volume to absorption, which devolves into a "reverberation time" of logarithmic decay of intensity of sound. These times, on the order of seconds, are generic but indeterminate functions of architecture. The critical time for human perception, the aural integration period, where perception of time (intervals) passes over to inverse time (frequencies), is two orders of magnitude shorter, and can be used to scale and define the "orchestral region" of a hall. In 25 ms sound travels 8.5 m, and 8.5 m is the wavelength of 40 Hz, near low E, the lowest orchestral note. If, by analogy with optics, 8.5 m is thought of as a "coherence length," it affects the time domain by promoting ensemble and the frequency domain by enhancing bass tones through proximity effects. A primary property of human hearing so scales the stages of concert halls.

9:00

5aMU3. Suppression of wolf-tone based on equivalent circuit model. Kei Ogura (Graduate School of Systems and Information Eng., Univ. of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan, ogura@aclab.esys.tsukuba.ac.jp), Koichi Mizutani, and Naoto Wakatsuki (Faculty of Eng., Information and Systems, Univ. of Tsukuba, Tsukuba, Japan)

Large size bowed stringed instruments, such as cello can pose undesired phenomenon called "wolf-tone" when you play at the specific pitch. When it occurs, a body strongly vibrates and a bow leaps on a string continuously. So, we are trying to figure out the cause of the wolf-tone and control it effectively. Thus, we devised "body-string coupled model" where body and string are coupled at bridge, and "bowed-string model" which expresses the relation between frictional force and relative speed, as it is called "stick and slip motion." We simulated using finite-difference time-domain method and figured out the result that reproduces particular string motion called "Helmholz motion." In addition, we set two parameters, speed of bow and pressure of bow, and calculated sound wave, oscillation mode, and the number of stick and slip. Then, we got the result of some existence ranges for wolf-tone. On this simulation, we added a small resonator in order to suppress the specific resonance mode which provides wolf-tone and observed wave forms. Based on these simulation results, we verified its effect experimentally on a real instrument.

5aMU4. Investigations of the coupling of doubled mandolin strings using high-speed video. Steve Tufte, Zach Rose, and Karlie Schwartzwald (Dept. of Phys., Lewis & Clark College, MSC 15, 0615 SW Palatine Hill Rd., Portland, OR 97219, tufte@lclark.edu)

The mandolin has four courses of doubled strings and in each pair the strings are tuned in unison. By using a macro lens on a quality high-speed camera, we are able to study the bridge-coupled motions of the strings in exquisite detail. We use an angled mirror to allow monitoring of the vertical and horizontal motions of both strings in a pair. With small colored tags on each string, we have automated the amplitude vs. time measurements using the Tracker Video Analysis tool. We observe clear evidence of coupling including the exchange of energy between the four modes. There is a strong tendency for the two strings to anti-correlate, while the vertical and horizontal modes of one string usually differ in phase by 90 degrees. We clearly observe the double-decay phenomena noted in Weinreich's study of piano strings. The coupling thus serves to add richness to mandolin tones and also increases the sustain. We compare the experimental results to a Runge-Kutta coupled oscillator model implemented in Matlab. Matching the observations requires the model damping constants to be slightly larger than the coupling constants and a slight detuning of the strings is necessary to produce the observed anti-correlated behavior.

5aMU5. Higher-order frequency locking of an organ pipe. Masahiro Okada and Tokihiko Kaburagi (Graduate School of Design, Kyushu Univ., 4-9-1 Shiobaru, Minami-ku, Fukuoka 815-8540, Japan, 3DS16004N@s.kyushu-u.ac.jp)

To investigate frequency locking of an organ pipe, Abel *et al.* forced a sounding pipe with a pure tone (J. Acoust. Soc. Am. **119**, 2467-2475, 2006). They showed that when the frequency of the pure tone approaches that of the pipe's fundamental, frequency locking can take place, whereby both frequencies are identical; that is, they are synchronized with a 1:1 frequency ratio. To gain a deeper understanding of organ pipes from the viewpoint of nonlinear oscillators, we searched for frequency locking at different ratios of the pure tone to the pipe's fundamental (pipe : pure tone = m:n). First, experiments were carried out for m:n = 1:2, 1:3, 2:1, and 2:3. We demonstrated that frequency locking occurs only if m = 1; that is, when the pure tone approaches the second or third harmonic of the pipe sound. We then analyzed our observations using synchronization theory, which describes frequency locking of a limit cycle oscillator. We found that m should be one when the external signal is a pure tone. The behavior of the organ pipe agrees well with that predicted according to synchronization theory; hence, the organ pipe can be treated as a nonlinear oscillator.

5aMU6. The effect of harmonic overtones in relation to "sharpness" for perceived brightness of distorted guitar timbre. Koji Tsumoto, Atsushi Marui, and Toru Kamekawa (Tokyo Univ. of the Arts, 1-25-1 Senju, Adachi-ku, Tokyo 120-0034, Japan, tsumoto@tcn-catv.ne.jp)

Perceived timbral brightness is often predicted by "sharpness," which is the conventional predictor described by von Bismarck. Perceived brightness and the predicted "sharpness" value increases when the energy of harmonic overtones in high-frequency range increases. Studies regarding the relationship between "sharpness" and harmonic overtones are relatively rare. "Sharpness" and harmonic overtones may have the synergistic effect or the canceling effect over perceived brightness. The electric guitar is one of the representative musical instruments which controls its timbre by adjusting the harmonic overtones. Under the assumption that there is a specific canceling effect, the stimuli generated by non-linear distortion processor for the electric guitar were compared for the first experiment. The values of the "sharpness" were adjusted by shelving filters. The onset of the each stimulus was deleted for the second experiment, since any variables other than "sharpness" and the harmonic overtones needed to be eliminated. The result indicated that "sharpness" and the amount of harmonic overtones had a canceling effect. A high amount of harmonic overtones decreased the perceived brightness. This result provokes the question of "why did the harmonic overtones cancel the perceived brightness predicted by sharpness?" This canceling effect was discussed by observing acoustic features.

5aMU7. Acoustic design of timpani using vibro-acoustic numerical analysis. Yozo Araki (Graduate School of Design, Kyushu Univ., 4-9-1 Shiobaru, Minami-ku, Fukuoka 815-8540, Japan, araki.yozo.105@s.kyushu-u.ac.jp) and Toshiya Samejima (Faculty of Design, Kyushu Univ., Fukuoka, Japan)

A vibro-acoustic analysis method of timpani reveals several conditions taking account of the interdependency among design parameters of timpani for delivering harmonic overtones. The head of a timpani is analyzed with the modal expansion type of the theoretical solution for an ideal circular membrane. The sound field around the timpani is analyzed with the normal derivative form of boundary element method. The theoretical solution of the membrane is included into the boundary element method as a characteristic matrix through proper vibro-acoustic coupling conditions. Therefore, the vibration displacement of the membrane is not included as unknowns in the equation to be solved. Frequency response functions of a timpani calculated with the analysis method show good agreement with measured ones. A timpani is designed based on the analysis method such that it has a strong sense of pitch. Tensions and surface densities of heads and kettle shapes that shift eigenfrequencies of the timpani into nearly harmonic ratios are presented. It is concluded that the consideration of the interdependency among such design parameters of timpani is important for composing harmonic overtones.

5aMU8. The effect of adaptive music playing system on emotion regulation. Wei-Chun Wang (Dept. of Humanities and Social Sci., National Taiwan Univ. of Sci. and Technol., No. 43, Sec. 4, Keelung Rd., Taipei 10607, Taiwan, vgnwang@hotmail.com)

We live in an industrially polluted environment, surrounded with stress and negative emotions. Since music has the marvelous power of arousing different moods and regulating emotions depending on the features of musical elements, people could be prescribed specific kinds of music to help alleviate their symptoms. Therefore, **refer to** *entrainment, iso, and diversion principles of music therapy stated by Campbell, the investigator developed an* Adaptive Music Playing System base on the two-dimensional arousal-valence emotion model. The purposes of this study were: (1) to explore the two-dimensional emotional responses of subjects when listening to specific music; (2) to survey the effects of gender, music preference on their perception of musical emotions; (3) to compare the relativities of various musical elements with musical emotions; (4) to detect the healing effect of adaptive music playing system. 56 college students were recruited as subjects in this study. Subjects imported their instant mood value, and the Adaptive Music Playing System would automatically play three music pieces. The results of this study showed that Adaptive music playing system could successfully lead the subjects in a better physical and mental condition. Gender effect wasn't significant for subjects' perception. Music preference affected subjects' perception.

5aMU9. Sixteen types' chord label estimation from acoustic signal of electric guitar. Nozomiko Yasui (National Inst. of Technol., Matsue College, 14-4 Nishi-ikuma-cho, Matsue, Shimane 690-8518, Japan, n_yasui@matsue-ct.jp) and Masanobu Miura (Faculty of Sci. and Technol., Ryukoku Univ., Otsu, Shiga, Japan)

It's difficult to determine an chord label for acoustical signal of musical playing, in particular when playing chords with "omitting", "inversions," or "tension voicing" on the guitar. Additionally, "enharmonic equivalence" produces multiple possibilities. This study developed a chord estimation system that deals with an audio signal output from electric guitars considering such techniques. All of the chord types employed in this study are the sixteen patterns frequently used in guitar chord playing. Chord labels are estimated by combination of salient pitch classes (or chroma), and some of them are dealt with as "performed notes" assumed as the member of played chord. Obtained performed notes are input to the "search tree for chord

labels" so as to search possible chord labels, by referring chord progression patterns included in the "chord progression database." Sixteen chord types are triad such as major, minor, aug, dim and sus4, with four-note chords such as 6th, 7th, Maj7, aug7, 7sus4, add9, min6, min7, minMaj7, min7(b5), and dim7. This study investigated appropriate threshold and type of filter used in the judgment of performed note. Results found that chord labels are estimated as 89% of accuracy when using three types of electric guitars.

11:00

5aMU10. Effect of music experience on perception of tempo change. Mayuko Yamashita, Masuzo Yanagida (Doshisha Univ., 1-3 Tatara Miyakodani, #221 Kochikan, Kyotanabe, Kyoto 610-0394, Japan, duq0171@mail4. doshisha.ac.jp), Ichiro Umata (KDDI R&D Labs., Tokyo, Japan), Tsuneo Kato, and Seiichi Yamamoto (Doshisha Univ., Kyotanabe, Kyoto, Japan)

Tempo is one of the basic factors in music expression and perception. Although there have been studies on the perception of tempo change, little is known about how the type of music experience affects the sensitivity to this change. We analyze the effects of music experience on the perception of tempo change to contribute to music education. Our analysis focuses on sensitivity to tempo change. Participants were classified into three groups according to their musical experience: (A) inexperienced in any musical instrument, (B) players majoring in the piano, and (C) amateur players belonging to brass bands. We performed experiments in which monotone piano sequences that gradually change tempo from the initial inter-onset interval (IOI) to the target IOI were used. We manipulated three tempo change patterns, namely, (I) linear, (II) exponential, and (III) the average of (I) and (II). We compared the point of tempo change perception among the three groups with the assumption that the sensitivity would be higher in (B) than in (A). The results, however, showed that the sensitivity was lowest in (B), except for one player who often plays with other players. These results suggest that ensemble experience may affect sensitivity to tempo change.

11:15

5aMU11. Automatic arrangement for ensemble music by estimating playing difficulty on instrumentalists. Nozomiko Yasui, Shiori Nabara (National Inst. of Technol., Matsue College, 14-4 Nishi-ikuma-cho, Matsue, Shimane 690-8518, Japan, n_yasui@matsue-ct.jp), and Masanobu Miura (Faculty of Sci. and Technol., Ryukoku Univ., Otsu, Shiga, Japan)

Since arrangement of sheet music for orchestral ensemble requires a certain musical knowledge and expertise, amateur orchestral musicians often have difficulties when arranging. An automatic arrangement method was proposed by employing Eigenmusic which is a set of eigenvector for many music excerpts, in order to evaluate the similarity among phrases. The difficulty in playing the arranged sheet music, however, was not evaluated; so, whether players feel difficulties on playing, or playability, was out of discussion. This study proposes a method to evaluate the degree how easy to play the phrases on sheet music. This method generates phrases for woodwind instruments using a MIDI database of backing phrases. Each musical phrases selected from the database are allocated to instrumentalists. In turn, the phrases' playability is evaluated based on instrumentalist's individual ratings for playing features such as cross-fingering, jumping notes, and so forth. Finally, phrases evaluated as the easiest are conjoined in each instrumentalist. We conducted an experiment to investigate the playability of generated sheet music by musician's rating. Experimental results showed that the evaluation of playability by proposed method is effective so that proposed method generates easier phrases to play than our previous method.

5a FRI. AM

Session 5aNSa

Noise: Wind Turbine Noise

Yasuaki Okada, Cochair
Meijo University, 1-501 Shiogamaguchi, Tempaku-ku, Nagoya 468-8502, Japan

Paul D. Schomer, Cochair
Schomer and Associates Inc., 2117 Robert Drive, Champaign, IL 61821

Nancy S. Timmerman, Cochair
Nancy S. Timmerman, P.E., 25 Upton Street, Boston, MA 02118

Invited Papers

8:00

5aNSa1. On translating minimum siting distances into percentages of receiving properties meeting a stated dB(A) criterion. Paul D. Schomer and Pranav K. Pamidighantam (Schomer and Assoc. Inc., 2117 Robert Dr., Champaign, IL 61821, schomer@SchomerAndAssociates.com)

Around the world, minimum siting distances are used by regulators and developers to limit the effects of wind turbine noise on people. Acousticians know that the proper calculation is equal sound level contours, but customers, in this case the communities, developers, and regulators, all want simpler solutions. This study makes use of data collected at over 1200 dwelling units as a part of the Health Canada Study. This paper provides a method to determine minimum siting distances based on predicted percentages of exceedances of dB(A) criterion at dwellings.

8:20

5aNSa2. Can inaudible and audible low level Infrasound and low frequency noise be an acoustic trigger of the startle reflex? Steven E. Cooper (The Acoust. Group, 22 Fred St., Lilyfield, NSW 2040, Australia, drnoise@acoustics.com.au)

Modulation of audible low frequency ventilation fan noise reproduced in a laboratory has been shown to trigger the startle reflex in people sensitised to that noise. Using a range of low frequency and infrasound noise signals to sensitised and unsensitized subjects can show a causal relationship between an acoustic trigger and a physiological stress response, which engineers call "annoyance" or "noise annoyance" symptoms, and which biologists recognize as the "startle reflex." Utilizing low level amplitude and frequency modulation as the source triggers a "startle reflex" response for comparison with the typical "startle reflex" to high level noise impulses.

8:40

5aNSa3. Wind farm infrasound—Are we measuring what is actually there of something else—(Part 3). Steven E. Cooper (The Acoust. Group, 22 Fred St., Lilyfield, NSW 2040, Australia, drnoise@acoustics.com.au)

With the method of analysis demonstrated in part 2 of the subject title, and the requirement for accurate reproduction of the signal (presented in Salt Lake City) the hypothesis as to infrasound contributions is revised in light of further investigations with test subjects for both inaudible ILFN and audible LFN to 50 dB(A). The work ties in with the investigations into the startle reflex using wind turbines as the source of the acoustic trigger. The ramifications as to suitable criteria for the protection of the community are discussed.

9:00

5aNSa4. Measures of sway as perceptual response to infrasound. Peggy B. Nelson, Michael Sullivan, Meredith Adams, and Andrew Byrne (Ctr. for Applied/Translational Sensory Sci., Univ. of Minnesota, 164 Pillsbury Dr. Se, Minneapolis, MN 55455, peggynelson@umn.edu)

Some reports suggest possible negative human perceptual responses to infrasound and audible noise from wind turbines. Recordings were made in the field of turbine noise (audible and infrasound) at 300 meters. In our multisensory laboratory (Center for Applied and Translational Sensory Science at the University of Minnesota) we are re-creating these natural stimuli in a controlled environment. Infrasound stimuli were presented at approximately 85 dB SPL; audible stimuli were 55 dB SPL. Stimulus configurations were combinations of audible and infrasound emissions, testing the effects of these stimuli individually and in combination. We measured sway using an AMTI AccuSway Optimized force plate. Preliminary data suggest that measures of sway are stable and reproducible, and that some persons may show changes in stance in the presence of combinations of audible and infrasound stimuli. [Funding provided by Xcel Energy RD-14 to the University of Minnesota.]

5aNSa5. Observations of vibration velocity on wind turbine with propagated sound to surroundings. Teruo Iwase (Niigata Univ., 8050 Igarashi 2 Nocho Nishi-ku, Niigata 950-2181, Japan, iwase@cc.niigata-u.ac.jp), Hideo Uchida (NS Environ. Sci. Consultant, Saitama, Japan), Hiroyasu Kurono (Faculty of Eng., Meijo Univ., Niigata, Japan), Yasuaki Okada, and Koichi Yoshihisa (Faculty of Sci. and Technol., Meijo Univ., Nagoya, Japan)

The authors newly tried measurements of the excited vibration on many parts of wind turbine such as outside surface of nacelle storing power generation system and tall tower by using a laser Doppler vibration meter. Observation and analysis of the natural vibration frequencies of blade in the stopping were done. Observations on sound in surroundings of wind turbine were also done. FFT analyses on them with high resolution to obtain detailed frequency characteristics and to know the relations between vibration velocity and propagated sound to surroundings were done in addition to ordinal spectral analyses. Sharp spectra at near 1 Hz as slightly lower or higher were appeared in the analysed results on blade and tower. In both the analyzed results of vibration velocity on wind turbine and propagated sound, a lot of sharp spectral peaks were recognized in wide frequency range from very low frequency to several hundreds of hertz with high coherencies between vibration velocity and propagated sound at their peak spectral frequencies. Certain values of coherency were remained even for the case of sound received in about 200 m distant from a wind turbine. These observed results would be effective for understanding key parts to make sound radiation lower.

5aNSa6. Experimental study on noise radiation from wind turbines. Yasuaki Okada, Yui Mizutani, Koichi Yoshihisa (Meijo Univ., 1-501 Shiogamaguchi, Tempaku-ku, Nagoya, Aichi 468-8502, Japan, okada@meijo-u.ac.jp), and Teruo Iwase (Niigata Univ., Niigata, Japan)

Noise emitted from wind turbines is composed of aerodynamic and mechanical sound and has directional radiation characteristics. To investigate the horizontal sound directivity around a wind turbine under various wind conditions, field measurements of noise generated from two different wind turbines have been performed over long periods. Wind turbine operational data such as the nacelle direction and rotor rotational speed were collected at 1 s intervals along with corresponding acoustic data. An empirical formula for the directivity correction was derived from the A-weighted sound pressure levels measured at the several receiving points around the wind turbine. We also focused on the amplitude modulation components of wind turbine noise in emission areas and compared the estimated rotor speeds by using measured sound pressure levels with actual values. The results showed that the directivity pattern of the A-weighted sound pressure level for two different wind turbines is almost the same, whereas the frequency dependence of the sound directivity is different for the individual wind turbines. Additionally, the radiation characteristics of wind turbine noise depend strongly on the rotor rotational speed, which can be estimated from the blade-passing-frequency detected by using the amplitude modulation components of wind turbine noise.

10:00–10:15 Break

Contributed Papers

5aNSa7. The influence of periodic wind turbine noise on infrasound array measurements. Christoph Pilger and Lars Ceranna (BGR Hannover, Stilleweg 2, Hannover 30655, Germany, christoph.pilger@bgr.de)

Aerodynamic noise from the continuously growing number of wind turbines in Germany creates increasing problems for infrasound array measurements recording acoustic signals at frequencies below 20 Hz. Ten years of continuous data (2006-2015) from the 4-element infrasound array IGADE in Northern Germany are analysed to quantify the influence of wind turbine noise in terms of enhanced amplitude modulations. Furthermore, a theoretical model is derived and validated by a field experiment with mobile microbarometers. Fieldwork was carried out to measure the infrasonic pressure level of a single horizontal-axis 200 kW wind turbine and to extrapolate the noise effect for turbines with higher electric powers and for a larger number of collocated wind turbines. The model estimates the generated sound pressure level of wind turbines and thus enables for specifying the minimum allowable distance between wind turbines and infrasound stations for undisturbed recording. This aspect is particularly important to guarantee the monitoring performance of the German infrasound stations I26DE in the Bavarian Forest and I27DE in Antarctica. These stations are part of the International Monitoring System (IMS) verifying compliance with the Comprehensive Nuclear-Test-Ban Treaty (CTBT), and thus have to meet stringent specifications with respect to infrasonic background noise.

5aNSa8. A numerical study of inflow turbulence distortion in the vicinity of blade leading edges. Thomas Hainaut, Gwenael Gabard, and Vincent Clair (Inst. of Sound and Vib. Res., Univ. of Southampton, Bldg. 13 - R2009, University Rd., Southampton SO171BJ, United Kingdom, t.hainaut@soton.ac.uk)

The interaction of the rotor with inflow turbulence is a source of broadband noise, dominant at high wind speeds. In the vicinity of the leading edge of realistic blades, the mean flow distorts the turbulence, resulting in an attenuation of the high-frequency part of the radiated noise compared to zero-thickness blades. In this paper, to study this turbulence distortion by the mean flow, two different numerical approaches are considered. First, the linearized Euler equations are solved in the time-domain using a finite difference code to model the response of an isolated blade interacting with synthetic turbulence. Second, a vorticity approach is applied, using the Biot-Savart law combined with a vortex panel method. Using these approaches on multiple configurations, the up-wash velocity fluctuations along a streamline which goes to the stagnation point, show a decrease of the turbulent levels, from a threshold distance independent of the wavenumber. This decay is then inverted for low frequencies after a wavenumber dependent threshold distance. The turbulence characteristics are found to have no effect on the distortion, whereas the geometry forward the position of the maximum thickness has an effect on this distortion. The full paper will contain more physical insights responsible for this distortion.

10:45–11:45 Panel Discussion

Session 5aNSb

Noise, ASA Committee on Standards, and Architectural Acoustics: Innovations in Floor Impact Noise Testing and Evaluation

David Lubman, Cochair
DL Acoustics, 14301 Middletown Ln., Westminster, CA 92683-4514

John LoVerde, Cochair
Veneklasen Associates, 1711 16th St., Santa Monica, CA 90404

Manabu Tanaka, Cochair
General Building Research Corporation of Japan, Fujishiro-dai 5-8-1, Suita, Osaka 5650873, Japan

Chair's Introduction—10:00

Invited Papers

10:05

5aNSb1. Acoustic characterization of in-room footfall noise. Samantha B. Rawlings (Veneklasen Assoc., 1711 16th St., Santa Monica, CA 90404, srawlings@veneklasen.com), Joshua A. Magee (Lawrence Livermore National Lab., Livermore, CA), Robert M. Tanen (Shen Milsom Wilke, New York City, NY), Jonathan C. Silver (none, Tel Aviv, Israel), and Robert D. Celmer (Acoust. Program & Lab, Univ. of Hartford, West Hartford, CT)

Existing literature concerning footfalls is primarily focused on its transmission between spaces, such as a floor/ceiling's Impact Isolation Class. Studies at the University of Hartford measured in-room sound power spectra produced by footfalls on twelve different floor surfaces using human subjects and a standard tapping machine. Within a qualified reverberation room, fourteen male and female subjects walked on the floor surfaces while wearing three different types of footwear: leather-soled shoes (hard), rubber-soled shoes (medium), and sneakers (soft). Sound power spectra and vibratory signatures were measured in 1/3 octaves according to the ISO 3741 standard. A tapping machine was also used on each floor profile using both standard drop weights and with cored samples of the same shoe soles attached to the bottom of each weight. The data for each floor profile produced averages by shoe type along with corresponding 95% confidence intervals. Correlations between sound power and vibratory spectra produced with human walkers versus tapping machine (both with and without "shoes") were investigated, resulting in correction factors to model machine tapping as human footfalls. [Work supported by The Paul S. Veneklasen Research Foundation.]

10:25

5aNSb2. American Society of Testing Materials draft standard for rating in-room floor impact noise. Jerry G. Lilly (JGL Acoust., Inc., 5266 NW Village Park Dr., Issaquah, WA 98027, jerry@jglacoustics.com)

In 2010 an ASTM subcommittee was formed to assess the possibility of rating the noisiness of different floor surfaces from impacts such as footfalls and dropped objects. Since that time the subcommittee has been meeting biannually to discuss potential impact noise sources and measurement techniques that could be used in a future standard. Different impact sources were evaluated including footfalls with different shoe types, golf balls, and the standard mechanical tapping machine. Both field and laboratory measurements were considered. This paper presents some of the preliminary test results that were obtained and the current status of the proposed laboratory standard.

10:45

5aNSb3. Lateral impact noise isolation: motivation, methods, and mitigation. John LoVerde and David W. Dong (Veneklasen Assoc., 1711 16th St., Santa Monica, CA 90404, jloverde@veneklasen.com)

Impact noise isolation measurements and evaluation are traditionally defined and considered for vertical adjacencies exclusively, where the impact source is located in the space above the receiving room. ASTM E1007 has specific protocols, procedures and methods for the measurements and analysis of the data collected. However, impact noise transmission into the receiving room laterally adjacent to the source room (i.e., on the same floor) is important in some conditions (e.g., condominiums, apartments, and hotels). The existing standard provides no guidance for performing a lateral impact noise isolation measurement. The authors have developed methods and practices for performing this measurement. These are presented along with a selection of results demonstrating potential issues, criteria, and mitigation.

5aNSb4. Difference in subjective magnitude between heavy-weight floor impact sources using auditory experiment. Jongkwan Ryu and Minju Jeong (Chonnam Univ., School of Architecture, Chonnam National University, 77 Yongbong-ro, Buk-gu, Gwangju 61186, South Korea, jkryu@jnu.ac.kr)

Difference in subjective magnitude between heavy-weight floor impact sources was investigated using two auditory experiments. In the first experiment, subjective magnitude to various level (41~56 dB, L'i,Fmax,Aw) of heavy-weight floor impact sound by bang machine, impact ball and children's jumping was investigated using rating scale method (7 points scale). And then, point of subjective equality (PSE) of floor impact sounds by bang, ball and children's jumping was drawn using relative annoyance judgement method in second experiment. Stimuli were pairs consisting of the reference (fixed sound level in 50 dB, L'i,Fmax,Aw) and test stimulus (varied sound level), and the subjects rated the relative annoyance of test stimuli compared with a reference stimulus using a numerical scale (-3~0~+3). Results showed that floor impact sound level by bang machine was 3~4 dB lower than that of impact sound by impact ball in the PSE. It was also found that subjective magnitude of floor impact sound by ball was quite similar as impact sound by children jumping.

Contributed Paper

5aNSb5. High-sound-insulation double floor structure with Helmholtz resonators: Experimental study on effect of resonator specifications. Yosuke Yasuda, Syunpei Hirose, Hidehisa Sekine (Dept. of Architecture, Faculty of Eng., Kanagawa Univ., Rokkakubashi 3-27-1, Kanagawa-ku, Yokohama-shi, Kanagawa 221-8686, Japan, yyasuda@kanagawa-u.ac.jp), and Mitsuru Yabushita (YAB Corp., Yokohama, Kanagawa, Japan)

We have developed a high-sound-insulation double floor structure using Helmholtz resonators. The main targets of this structure are a lot of old apartment buildings that should be renovated, which have thin slabs and cannot sufficiently insulate floor impact sound. The proposed double floor structure greatly improved the insulation performance for heavy-weight floor impact sound in a real old apartment building with 110-mm slab: about 10 dB for 63-Hz and 125-Hz octave bands. In this presentation, experimental results are shown using a small unit structure partially taken out from the whole double floor structure to thoroughly grasp the actual vibration phenomenon. The results are compared to calculation results obtained from a two-particle system theory. Main conclusions are summarized as follows: (i) A sufficiently large rigidity of the resonator walls and a large number of resonator necks are required for obtaining a clear dip in the vibration transmissibility as calculated from the two-particle system theory. (ii) The horizontal and vertical displacement of the resonator necks hardly affects the vibration transmissibility. (iii) The effects of other factors such as the neck length and neck number are qualitatively predictable by the two-particle system theory.

Session 5aPA

Physical Acoustics: Ultrasonics and Non-Destructive Testing

Jinying Zhu, Chair

Civil Engineering, University of Nebraska-Lincoln, 1110 S 67th St, PKI 204C, Omaha, NE 68182

Contributed Papers

8:30

5aPA1. A non-contact detection of the internal defect in solid material by aerial ultrasonic beam. Yuriko Mukaiyama (Elec. Eng., Nihon Univ., 4-34-11, Ohizumimachi, Nerima-ku, Tokyo-to 178-0062, Japan, csyu16035@g.nihon-u.ac.jp), Ayumu Osumi, and Youichi Ito (Elec. Eng., Nihon Univ., Chiyoda-ku, Tokyo-to, Japan)

Measurement method using aerial sound waves and an optical equipment is utilized in a non-contact and non-destructive testing. Technique that we have been conventionally proposed, uses high-intensity aerial ultrasonic waves (20kHz). Specifically, a surface of an object is vibrated by high-intensity ultrasonic waves with a point focus. A vibration generated on the surface of the object is measured by laser Doppler vibration galvanometer. In general, the vibration of defect area is larger than that of defect-free area when the sound waves were irradiated to the object. Therefore, by performing the measurement over a target area, it is possible to image the defect area from the vibration distribution. However, this method needs a long time to measure the object with the defect. Therefore, we adopted a method of irradiating high-intensity ultrasonic to a wide range on the surface of the object using an aerial ultrasonic waves focused into a beam. In this report, we describe the details of the experimental results to detect defects in the object by using this technique.

8:45

5aPA2. Non-contact method for detecting crack in shallow layer of solid materials by using very high-intensity aerial ultrasonic wave and optical equipment. Ayumu Osumi (Nihon Univ., 1-8,KandaSurugadai, Chiyoda 101-8308, Japan, oosumi@ele.cst.nihon-u.ac.jp), Masashi Ogita, and Youichi Ito (Nihon Univ., Chiyoda-ku, Japan)

We have studied a noncontact and nondestructive method of imaging by using a high-intensity aerial ultrasonic waves and an optical equipment. This non-contact method detects a defect in solid materials by measuring a vibration velocity distribution on an object surface continuously irradiated by the aerial ultrasonic waves. In previous study, it was confirmed to image a relatively large peeling and defect in size by proposed method. Meanwhile, one of nondestructive testing target is a crack. It is difficult to vibrate the crack by this method has a lack of sound pressure of irradiating sound wave because the crack is very smaller than peeling. Therefore, in this study, we developed a new very high-power aerial ultrasonic sound source and attempted to detect the crack. In this report, we investigated to detect a crack in shallow layer of solid materials by this ultrasonic sound source. As a result, it is confirmed to image the crack around the shallow layer of object and the micro crack for extending to front surface of object by proposed method.

9:00

5aPA3. Acoustic resonances greatly improve detection of leaks from large pressure vessels. Keith A. Gillis and Michael R. Moldover (Sensor Sci. Div., National Inst. of Standards and Technol., 100 Bureau Dr., Mailstop 8360, Gaithersburg, MD 20899-8360, keith.gillis@nist.gov)

We measured the pressure p and the acoustic resonance frequency f_a of argon gas in a 300 L tank. We then used the ratio p/f_a^2 (proportional to the

mass M of gas) to determine the fractional leak rate $(dM/dt)/M \approx -13 \times 10^{-6}$/hr of gas from the tank. The tank was filled with argon at 450 kPa and was exposed to sunshine-driven temperature and pressure fluctuations as large as $(dT/dt)/T \approx (dp/dt)/p \approx 5 \times 10^{-2}$ h^{-1} (i.e., $1000 \times$ larger) in a 24-hour record. This leak could not be detected in a 72-hour record of p/T. Theory predicts that the resonance frequencies are insensitive to linear temperature gradients; however, we do not have a detailed understanding of why the acoustic frequencies are insensitive to asymmetric, time-dependent heating of the tank. From auxiliary measurements and information from the literature, we believe that convective currents near the tank's walls carry most of the heat between the hot and cool regions, leaving a linear temperature gradient in the majority of the gas's volume. We plan to report preliminary measurements for a much larger (1800 L) tank. We expect that the results of this investigation will lead to effective acoustic methods for flow calibrations and leak detection in large volumes.

9:15

5aPA4. Bonding characterization of three layers metal/adhesive/metal using Lamb waves. Camille Gauthier (LOMC, Univ. of Le Havre, France, Le Havre, France), Mihai Predoï (Univ. Politechnica of Bucarest, Bucarest, Romania), Ech-Cherif El-Kettani Mounsif (LOMC, Univ. of Le Havre, France, IUT, Pl. Robert Schuman, B.P. 4006, Le Havre 76610, France, elkettani@univ-lehavre.fr), Jocelyne Galy (IMP INSA LYON, Lyon, France), Damien Leduc, and Jean-Louis Izbicki (LOMC, Univ. of Le Havre, France, LE HAVRE, France)

The studied samples are three layer Aluminum/Epoxy/Aluminum structures, realized with the cooperation of Physico-Chemists, to obtain samples of controlled adhesion properties, and acousticians for ultrasounds characterization. Different mechanical (roughness) and/or chemical (silanisation) surface treatments at the interface substrate/adhesive are performed. A numerical FEM model is achieved to solve a rheological description of the interface as springs distribution. This model allows the study of the sensitivity of Lamb waves to the interface and the determination of their energy in the cross section of the structure, depending on the stiffness values of the springs. The sensitive modes are then looked for to be excited in the experimental study, particularly the vertical modes that are indicative of a more or less coupling in the three layer structure. Experimental results show different adhesion levels, depending on the surface treatments, and correspond stiffness springs value to each sample.

9:30

5aPA5. Modeling the scattering of elastic waves from defects in buried pipes. Ray Kirby (Mech. Eng., Brunel Univ., Uxbridge, Middlesex UB8 3PH, United Kingdom, ray.kirby@brunel.ac.uk) and Wenbo Duan (Mech. Eng., Brunel Univ., London, United Kingdom)

Long Range Ultrasonic Testing (LRUT) is a popular non-destructive evaluation technique for identifying defects in pipelines. The method sends an elastic wave down the walls of a pipe and then monitors echoes that arise if the wave is scattered by a defect. The technical challenge of LRUT lies in separating out, and interpreting from coherent and random background noise, those signals that belong to a defect. This becomes particularly challenging when a pipe is buried, because energy in an elastic wave is known

3386 J. Acoust. Soc. Am., Vol. 140, No. 4, Pt. 2, October 2016

5th Joint Meeting ASA/ASJ **3386**

to leak out of the pipe walls when it is surrounded by materials such as soil, concrete, or sand. To address this problem it is necessary to develop a better understanding of how elastic waves propagate in buried structures, and so, a finite element based model is introduced here that seeks to analyze the scattering from a defect in a buried pipe in both the frequency and time domain. Results from the implementation of this numerical model for the torsional T(0,1) mode are presented, and the effects of the burying material on the LRUT time domain response are discussed.

9:45–10:00 Break

10:00

5aPA6. Characterization of a local scatterer in a plate from reverberation of flexural waves. Emmanuel Moulin (IEMN, UMR CNRS 8520, Univ. of Valenciennes, Le Mont Houy, Valenciennes F-59313, France, emmanuel.moulin@univ-valenciennes.fr), Hossep Achdjian (GREMAN, INSA Ctr. Val de Loire, Université François Rabelais, Blois, France), Farouk Benmeddour, Lynda Chehami, Jamal Assaad, and Lucie Dupont (IEMN, UMR CNRS 8520, Univ. of Valenciennes, Valenciennes, France)

In media where acoustic waves are subject to multiple propagation paths (scattering or reverberation), recorded acoustic signals present a random aspect. Still, such signals carry some information about the medium, and ensemble averaging can give access to the estimation of a number of useful parameters. A well-known example, in room acoustics, is the estimation of absorption coefficients of walls from average decrease of the reverberation envelopes. In the work reported here, we have proposed a statistical model allowing to relate the scattering properties of a local heterogeneity (defect) to the average properties of reverberated acoustic signals in a solid plate. A theoretical expression of the averaged envelope of signals produced by scattered and reverberated flexural waves has been derived and both numerically and experimentally validated. A simple curve-fitting procedure applied to signals recorded on a few receivers then allows an experimental estimation of the scattering cross-section of the heterogeneity.

10:15

5aPA7. Integrity evaluation of adhesive anchors using electromagnetic acoustic transducer. Kazuhiko Hasebe, Yosuke Mizuno, and Kentaro Nakamura (Lab. for Future Interdisciplinary Res. of Sci. and Technol., Tokyo Inst. of Technol., 4159-R2-26 Nagatsuta-cho, Midori-ku, Kanagawa-ken 226-8503, Japan, khasebe@sonic.pi.titech.ac.jp)

The aging of social infrastructures has recently become a serious problem. The high efficiency of the testing is required to handle increased inspection demand. There is interest in testing the soundness of adhesive anchors. Bolts are connected to a concrete using holes filled with adhesive materials. Since anchor bolts is inspected by skilled workers manually, it is difficult to improve the efficiency. Several sensing methods have been proposed, but these measurements require a precision alignment or contact. We propose a non-contact evaluation method of adhesive anchors using an electromagnetic acoustic transducer (EMAT). An EMAT is a transducer for non-contact vibration generation and reception based on electromagnetic mechanisms. A neodymium ring magnet and two coils were used as an EMAT. We used 100 mm long M8 and M12 steel bolts mounted with epoxy resin in 70 mm depth holes in concrete. The EMAT which was placed around the bolt excited longitudinal vibrations of bolt by burst waves. Vibration signals were received after the excitation. The relationships between the resin amount and the resonance characteristics of the vibrations are evaluated for pursuing the applicability of this method to the on-site inspection. Resonance frequencies and quality factors are utilized for health monitoring of adhesive anchors.

10:30

5aPA8. Liquid loading characteristics of multiple SH-wave roundtrip in c-axis parallel oriented ZnO film/silica glass pipe. Shinji Takayanagi (Nagoya Inst. of Technol., Gokiso-cho, Syowa-ku, Nagoya 466-8555, Japan, takayanagi.shinji@nitech.ac.jp), Shoko Hiyama, Mami Matsukawa (Doshisha Univ., Kyotanabe, Japan), and Takahiko Yanagitani (Waseda Univ., Shinjyuku, Japan)

Shear waves with in-plane displacement are suitable for liquid sensors because they can propagate without the energy leakage into liquid. In previous study, we have demonstrated the SH-SAW excitation with (11-20) oriented ZnO film whose c-axis is parallel to the substrate plane [1]. For a high sensitive sensor, the long propagation path is needed to detect small changes of the velocity and amplitude. Multiple wave roundtrips on a pipe surface realizes the long propagation path. In this study, we fabricated IDT/c-axis parallel oriented ZnO film/silica glass pipe structure. The fourth roundtrip was observed by the time response measurement using a network analyzer. The insertion loss of the first lap consisted of two frequency components at 131 MHz and 160-350 MHz. We also measured the liquid loading characteristics by immersing the inside pipe surface. Compared with the unloaded characteristics, insertion losses with pure water increased 0.6 dB at 131 MHz and 1.1 dB at 236 MHz. Because of small increases in the insertion losses, we can conclude that SH-type plate wave was excited. Viscosity measurements with the inside pipe surface are expected. [1] A. Tanaka, *et al.*, *IEEE TUFFC*, **55**, 2709 (2008).

10:45

5aPA9. Transmission and irradiation of high-intensity aerial focused ultrasonic waves using pipe. Norifumi Suzuki (Elec. Eng., Nihon Univ., 2-2-17, Imai, Chuo-ku, Dia Palace Keiyo Soga II #1205, Chiba-shi, Chiba-ken 260-0834, Japan, csno16021@g.nihon-u.ac.jp), Ayumu Osumi, and Youichi Ito (Elec. Eng., Nihon Univ., Chiyoda-ku, Tokyo-to, Japan)

In this study, we have developed a method that high-intensity aerial ultrasonic waves (sound pressure level: 180 dB) is transmitted to a far place and irradiate on optional and specified place without attenuation. To irradiate on the place that is optional and specified with high-intensity aerial ultrasonic waves, it is necessary to move the sound source. If any obstacle is placed between the specified place and the sound source, however, it will be difficult to irradiate to that place. In addition, there will also be a problem that ultrasonic waves attenuate during propagation in air. If these problems can be solved, it will be expected that more effective irradiation with ultrasonic waves may be realized and the range of applications may be widened. To solve the problem above, we proposed to use a rigid wall pipe whose inner diameter is similar to a wavelength to transmit the sound waves and attempted to simulate and experiment on the problem. In this report, we confirmed that the high-intensity aerial focused ultrasonic waves (frequency: 20 kHz) can be transmitted by using the pipe.

11:00

5aPA10. Non-contact heating of surface on the object using by high-intensity aerial ultrasonic waves. Hirotaka Sato (Elec. Eng., Nihon Univ., 2-2-1, Omotechou, Sakura-shi, Chiba-ken 285-0811, Japan, cshi15015@g.nihon-u.ac.jp), Ayumu Osumi, and Youichi Ito (Elec. Eng., Nihon Univ., Chiyoda-ku, Tokyo-to, Japan)

In the present study, the authors investigate a method of rising a temperature of an object's surface for an instant without contact by using high-intensity aerial ultrasonic waves (frequency 20 to 50 kHz). In recent years, much attention has been directed at a thermal effect technology which utilizes ultrasonic energy such as HIFU and MRgFUS treatment. These technologies are realized for heating by using ultrasonic source which contact with a target. It is expected to realize the same heat effect by aerial ultrasonic waves, but a research of using ultrasonic waves to heat objects without contact has never been done. The authors propose a new method to use a very high-intensity aerial ultrasonic waves (intensity is about 25 kPa) which are generated by a point-converging acoustic source with a stripe-mode vibration plate and attempt to establish a technique. Recently, the authors reported that a temperature of silicone rubber material rise about 15°C after 1 second from the start of ultrasonic irradiation. With the realization of the technology, it is able to be a new underlying technology and is expected to be applied to a development of heating technique and non-destructive inspection.

11:15

5aPA11. A novel high-volume separation technology base on acoustophoresis for treatment of produced water in the oil and gas industry. Jason P. Dionne (FloDesign Sonics, Inc., 499 Bushy Hill Rd., Simsbury, CT 06070, j.dionne@fdsonics.com) and Bart Lipkens (FloDesign Sonics, Inc., Wilbraham, MA)

A novel emulsion splitting separation technology has been developed using acoustic radiation forces to recover the dispersed oil phase from an oil water emulsion. Current separation technologies suffer from high energy costs, use of consumables, fouling, and limited efficiency in separation of micron-sized particles. Multi-dimensional ultrasonic standing waves are used to trap a dispersed phase in a fluid. The action of the acoustic forces results in clustering and coalescence of droplets. Upon reaching a critical size, they are continuously separated through enhanced buoyancy. A second mode of operation uses acoustic radiation forces to increase the average droplet size and reduce the sub-20 micron droplet concentration. Earlier work was presented at ICA 2013(Dionne [1]). New results are shown for prototypes with a 1x2, 3x4 and 6x6 inch flow chamber driven by 2 & 3 MHz PZT transducers operating at flowrates of 1L/h, 30L/h, and 227-1136 L/h, tested with produced water samples from four US locations. Measured separation efficiencies of over 90% have been documented as well as particle size shifts of >100 micron. The particle size shift is particularly appealing as it acts as a complimentary technology to existing oil & gas separation technology. [1] J. Dionne, B. McCarthy, B. Ross-Johnsrud, L. Masi, and B. Lipkens, "Large volume flow rate acoustophoretic phase separator for oil water emulsion splitting," J. Acoust. Soc. Am., Vol. **133**, No. 5, Pt. 2, May 2013, pp. 3237. [Work supported by NSF SBIR IIP- 1330287.]

11:30

5aPA12. Injection of liquid into ultrasonic standing wave fields by exciting flexural vibrations on needle. Haruna Tadakoshi, Hiroki Tanaka, Yosuke Mizuno, and Kentaro Nakamura (Lab. for Future Interdisciplinary Res. of Sci. and Technol., Tokyo Inst. of Technol., 4259-R2-26 Nagatuta, Midoriku, Yokohama, Kanagawa 226-8503, Japan, htadakoshi@sonic.pi. titech.ac.jp)

There is a large need to handle droplets without contact in the development of new drug and materials. We have been trying to utilize ultrasonic levitation for this aim. It is needed to inject a droplet of precise volume into required location of the levitation device. But it is difficult to control the droplet volume using a syringe due to the influence of surface tension. In this study, by exciting ultrasonic vibrations on a needle, we investigated the control method of droplet volume and ejection timing. Exciting flexural vibrations on a needle for 0.2 seconds, a water droplet of 18-28 mg was dropped with a good controllability and the timing of inject can be regulated. To elucidate this dropping mechanism, we observed the behavior of the droplet. Capillary waves were excited on the droplet, and the force of 0.7 times of gravity exerted by vibration was applied on the droplet according to our estimation. Next, we examined the effect of droplet volume on the resonance frequency of flexural vibrations of the needle, and we confirmed that there was a little influence between them. This means that there is no need to change the frequency in accordance with the volume of droplets.

11:45

5aPA13. Behavior of liquid in a thin type vessel irradiated with high-intensity aerial ultrasonic waves. Taichi Urakami (Elec. Eng., Nihon Univ., 1336-7, Nakao, Midori-ku, Saitama-shi, Saitama-ken 336-0932, Japan, cstt15003@g.nihon-u.ac.jp), Ayumu Osumi, and Youichi Ito (Elec. Eng., Nihon Univ., Chiyoda-ku, Tokyo-to, Japan)

A high-intensity aerial ultrasonic waves can remove liquid that has penetrated into the pores instantaneously. Further, tiny particles adhered to a solid surface is removed from the surface of object by radiation of the ultrasonic waves instantaneously. These techniques are implemented to a target by an action aerial ultrasonic waves directly. If these operations can be realized similarly even when the ultrasonic waves pass through the object, the field of application is expected to spread extremely. For example, if it is possible to exert a force to the liquid injected into a small container from the outside of the container by irradiating the aerial ultrasonic waves, various applications can be greatly expected. We have investigated the observations of liquid behavior when the high-intensity aerial ultrasonic waves were irradiated to the very small vessel filled with liquid. As a result, tiny bubbles were generated from the vicinity irradiated with the aerial ultrasonic waves, and the bubbles had a fierce tiny vibration. Further, in the observation of the liquid in the vessel by using ink, the ink is greatly diffused throughout the vessel and the behavior of the bubbles was confirmed to contribute to the diffusion of the ink extensively.

FRIDAY MORNING, 2 DECEMBER 2016   CORAL 5, 7:55 A.M. TO 12:00 NOON

Session 5aPP

Psychological and Physiological Acoustics: Pitch and Timbre

Xin Luo, Chair
Speech and Hearing Science, Arizona State University, 975 S. Myrtle Ave., Tempe, AZ 85287

Contributed Papers

7:55

5aPP1. Pitch model as interspike interval periodicity present at all active Bark bands during harmonic sound perception. Rolf Bader (Inst. of Systematic Musicology, Univ. of Haburg, Neue Rabenstr. 13, Hamburg 20354, Germany, R_Bader@t-online.de)

A Finite-Difference Time Domain (FDTD) physical model of the basilar membrane (BM) is implemented including the transition of mechanical energy on the BM into spike excitation. The spike train caused by this transition shows energy at all Bark bands of a harmonic input sound. Still higher partials are often not able to produce a regular interspike interval (IS) firing pattern all through one periodicity of the fundamental frequency. Instead drop-outs occur, where one or several spikes are missing. This produces higher ISI periodicities and therefore lower frequencies as the critical frequency of the respective Bark bands. These lower frequencies are integer subharmonics of the critical frequency. The lowest of these frequencies is the fundamental frequency of the input sound. Therefore, the model shows these fundamental frequencies in all Bark banks of active sound input

frequency. The fundamental frequency of a harmonic sound is therefore present at many cochlea fibers and may be the reason for pitch perception. This mechanism also works for residual pitches. Compared to other pitch models which need further calculation in following neural nuclei, like autocorrelations, the pitch model proposed here has pitch intrinsically present right at the cochlea output with no need of further computations. From this view, a pitchness can be defined as the strength of a frequency present at many Bark bands, in accordance with the strength of a pitch perception.

8:10

5aPP2. More about monaural noise edge pitch. William M. Hartmann and Aimee Shore (Phys. and Astronomy, Michigan State Univ., 567 Wilson Rd., East Lansing, MI 48824, hartmann@pa.msu.edu)

Lowpass noise with a sharp spectral cutoff generates a monaural noise edge pitch (MNEP) that listeners match with a sine tone having a frequency somewhat below the noise cutoff frequency (negative pitch shift). If the MNEP is a proper musical pitch, listeners should be able to identify musical intervals made using only the MNEP stimulus. Open-set MNEP experiments in the edge-frequency range of 600-2400 Hz found that four out of four listeners with musical training correctly identified octaves, fifths, fourths, and major thirds, proving that the MNEP is a musical pitch. The MNEP can be heard for edge frequencies at least as low as 100 Hz, but the negative pitch shift becomes larger than a semitone, consistent with a timing model of the MNEP having a fixed integration time. If the MNEP reflects the expected quasi-periodic timing response of the auditory system to a sharp spectral edge, the pitch should become inaudible if the edge frequency becomes too high for neural synchrony—expected to be about 5000 Hz. Sine tone pitch matching experiments showed that such a limit clearly applies for most listeners, but less clearly for others. [Work supported by the USAFOSR.]

8:25

5aPP3. The effect of pitch change direction on melodic interval ranking. Xin Luo (Speech and Hearing Sci., Arizona State Univ., 975 S. Myrtle Ave., Tempe, AZ 85287, xinluo@asu.edu)

Our previous study showed that listeners had significantly lower ascending interval ranking thresholds than descending ones. Experiment 1 tested whether the effect on interval ranking of local pitch change direction within intervals relies on the global pitch change direction between intervals. Both ascending and descending intervals were ranked in both the 2nd-interval-higher and lower conditions. Ascending interval ranking thresholds were significantly lower than descending ones in the 2nd-interval-higher conditions, but similar to or higher than descending ones in the 2nd-interval-lower conditions. Interval ranking was distracted when the global pitch change between intervals and local pitch change within intervals were in the opposite direction. To better simulate music listening, Experiment 2 measured ascending and descending interval ranking in the 3-tone conditions, where the three tones forming two intervals were played back to back in time. The 4-tone conditions with a 1-sec temporal gap between intervals were also tested. Similar ascending and descending interval ranking thresholds were observed when the global pitch change between intervals and local pitch change within intervals had the same direction. Interval ranking thresholds were significantly lower in the 4-tone conditions than in the 3-tone conditions, indicating that interval ranking was enhanced by the temporal gap between intervals.

8:40

5aPP4. Amplitude modulation masking for frequency modulation detection. Andrew J. King (Laboratoire des systemes perceptifs, CNRS, Ecole Normale Superieure, Paris Sci. et Lettres Res. Univ.,UMR CNRS 8248, Departement d'Etudes Cognitives, Ecole normale superieure, 29 rue d'Ulm, Paris 75005, France, andrew.king@ens.fr) and Christian Lorenzi (Laboratoire des systemes perceptifs, CNRS, Ecole normale superieure, Paris Sci. et Lettres Res. Univ., Paris, Ile de France, France)

It is generally assumed that frequency modulation (FM) is detected as amplitude modulation (AM) for fast FM rates with low carrier frequencies,

and for all FM rates with high carrier frequencies. If this is the case, then FM detection in the presence of an AM masker should exhibit the main features of AM masking: tuning, dependency on AM masker depth, negative masking, beating effects and phase effects. We explored the masking effects produced by sinusoidal AM on detection thresholds of sinusoidal FM for normal-hearing listeners. FM rates ranged between 2 and 64 Hz. The carrier was either a 500-Hz or a 5000-Hz pure tone that was either unmodulated in amplitude or modulated in amplitude at 2 or 16 Hz. The masker AM depth was fixed to either 50% or 25%, and stimulus duration was set to either 500 ms or 1 sec. Additionally, detection thresholds were tested as a function of the phase relationship between a 2-Hz FM target and a 2- or 4-Hz AM masker, and between a 16-Hz FM target and an 8-, 16-, or 32-Hz AM masker. The data will be discussed in light of previous studies on AM masking and the modulation filter-bank concept.

8:55

5aPP5. Effect of harmonic structure of notes on categorical chord perception. Chiaki Utsugi, Mariko Hamamura, and Kiyoaki Aikawa (School of Media Sci., Tokyo Univ. of Technol., 1404-1 Katakuracho, Hachioji, Tokyo 192-0982, Japan, mariko@hamamura.biz)

This paper reports distinct differences among the harmonic structures of notes on chord perception. Experimental results clarified that the existence of harmonics greatly improved chord perception accuracy. Subjective tests were conducted for 20 combinations of three notes from a minor chord to a major chord. The fundamental notes and the highest notes were fixed. The fundamental frequency of the middle note was shifted at an equal interval on the log scale. The harmonic structure was based on saw-tooth wave. The categorical perception accuracies were compared between pure tone and complex tone including five harmonics. The fundamental notes were C5 (523.3 Hz), A5 (880 Hz), and C7 (2093 Hz). The stimuli were diotically presented at 60 dB(SPL) through electro-static headphones. The subjects were requested to answer major or minor with the 2AFC method. The categorical perception accuracy was extremely low in case of C7 without harmonics. The percent corrects were as low as 46.7% and 54.8% for Cm and C, respectively. The Cm and C percent accuracies with harmonics were as high as 77.3% and 79.5%. The Cm and C accuracies were 97.7% and 93.2% for C5. The perceived categories turned from minor to major in the region between 8th and 15th note combinations.

9:10

5aPP6. Every advantage has its disadvantage: Pitch perception of tone language speakers is better at low frequencies but poorer at high frequencies. Yu-Xuan Zhang, Xiang Gao, Dinglan Tang, Ting Huang (National Key Lab. of Cognit. Neurosci. and Learning, Beijing Normal Univ., Rm. 330, Yingdong Bldg., Beijing 100875, China, zhangyuxuan@bnu.edu.cn), and Chang Liu (Dept. Commun. Sci. & Disord., Univ. Texas Austin, Austin, TX)

Tone language experience has been associated with more accurate and stronger phase locking of neural responses in the brainstem to auditory stimuli. However, whether this neurophysiological advantage translates to better pitch perception remains under debate. Previous studies have yielded controversial results, possibly due to variations in methods and large individual variability. Here we examined tone frequency discrimination (FD) over the frequency range of 250 Hz to 8 kHz in Chinese and English speakers using relatively large sample sizes (N>50). For both brief (7.5-ms) and steady (100-ms) tones, FD threshold showed an interaction between group and frequency: Chinese speakers outperformed English speakers up to 4 kHz, and the advantage was reversed at higher frequencies. The only exception was the 7.5-ms 250-Hz condition, where the stimuli were too short (< 2 cycles) for pitch extraction by phase locking and performance was worse in Chinese than English speakers. Overall, the pattern of FD differences between Chinese and English speakers indicates that tone language experience enhances temporal coding of pitch at the cost of place coding.

9:25

5aPP7. Training F0 discrimination improves vowel identification among speech m. Xiang Gao and Yuxuan Zhang (National Key Lab of Cognit. Neurosci. and Learning, Beijing Normal Univ., Yingdong Bldg. 327, No. 19 Xinjiekouwai St., Haidian District, Beijing, Beijing 100875, China, gaoxiang8945@163.com)

Listeners can use voice or fundamental frequency (F0) differences to separate target speech from competing talkers. Discrimination of F0 of harmonic tones has been shown to improve with adaptive training. Here, we examine whether such F0 learning can benefit speech identification in noise. Normal-hearing adults practiced F0 discrimination with standard F0 roved between 120 and 240 Hz. Before and after training, these listeners and a no-contact control group were tested on a vowel identification task, in which Mandarin vowels were embedded within six-talker babble noise at multiple signal-to-noise ratios. After training, the trained listeners improved more than controls on F0 discrimination. Critically, they also improved more in vowel identification. Threshold SNR at 60% intelligibility dropped 1.9 dB as a result of F0 training. The results demonstrate that speech perception in noise can benefit from learning of basic auditory skills, revealing a promising approach for rehabilitative training.

9:40

5aPP8. Development of music-specific pitch processing. Christina Vanden Bosch der Nederlanden, Erin Hannon, and Joel Snyder (Psych., Univ. of Nevada, Las Vegas, 4505 Maryland Parkway #455030, Las Vegas, NV 89154, erin.hannon@unlv.edu)

Both music and speech contain meaningful variations in pitch, however discrete pitch intervals are more prominent in music than they are in speech. Most research has separately examined perception of pitch patterning in the musical and linguistic domains; however, less is known about how listeners flexibly recruit musical and linguistic knowledge to interpret similar pitch patterns across different contexts, nor how these abilities develop with age. We asked children and adults to detect pitch changes to vocal stimuli that were presented in separate blocks as music (part of a "musical") or speech (part of a "play"). Stimuli were overtly sung (in the musical) or spoken (in the play) but matched in pitch range and contour, and across both contexts we occasionally presented ambiguous vocalizations (speech-to-song illusions). Pitch changes were created by subjecting stimuli to a pitch range expansion which preserved the pitch contour but distorted pitch interval relationships. Adults showed greater detection accuracy for overtly sung than spoken utterances; however, children showed comparable performance for speech and song. No difference was observed for ambiguous vocalizations at any age. Implications for the development of domain-specificity of pitch processing will be discussed.

9:55–10:15 Break

10:15

5aPP9. The relationship between perceived pleasantness and memory for environmental sounds. Kelly Dickerson (Army Res. Lab., 131 Waldon Rd., Abingdon, MD 21009, dickersonkelly23@gmail.com), Laura Sherry (Audiol., Speech-Lang. Pathol. and Deaf Studies, Towson Univ., Towson, MD), and Jeremy Gaston (Army Res. Lab., Aberdeen Proving Ground, MD)

The everyday auditory environment contains many sounds that can be influenced by semantic information and listener experiences. Some sounds are more memorable than others and there is limited work examining the role of specific semantic or subjective attributes on retrieval success for environmental sounds. The present study examines the relationship between pleasantness ratings and memory performance on a cued-recall task. A group of listeners were presented with 36 environmental sounds and asked to rate each for perceived pleasantness to provide baseline evaluations for each sound. Based on the ratings, stimuli were sorted into two pleasantness-balanced groups. A second group of listeners evaluated half of the stimuli for pleasantness during the "study" phase, which was followed by a cued recall test. At test, half of the items were "old" (from the study phase), and half were "new" (from the set of 36 sounds). The results demonstrate that listener's recalled more pleasant sounds more accurately than unpleasant

sounds. Further, listeners produced more false alarms for the more pleasant sounds. These results suggest that listeners are biased to report recalling more pleasant sound regardless if they had been heard previously.

10:30

5aPP10. Comparing identification of vocal imitations and computational sketches of everyday sounds. Guillaume Lemaitre, Olivier Houix, Frédéric Voisin, Nicolas Misdariis, and Patrick Susini (IRCAM, 1 Pl. stravinsky, Paris 75004, France, GuillaumeJLemaitre@gmail.com)

Sounds are notably difficult to describe. It is thus not surprising that human speakers often use many imitative vocalizations to communicate about sounds. In practice, vocal imitations of non-speech everyday sounds (e.g., the sound of a car passing by) are very effective: listeners identify sounds better with vocal imitations than with verbal descriptions, despite the fact that vocal imitations are often inaccurate, constrained by the human vocal apparatus. The present study investigated the semantic representations evoked by vocal imitations by experimentally quantifying how well listeners could match sounds to category labels. It compared two different types of sounds: human vocal imitations, and computational auditory sketches (created by algorithmic computations), both based on easily identifiable sounds (sounds of human actions and manufactured products). The results show that performance with the best vocal imitations was similar to the best auditory sketches for most categories of sounds. More detailed analyses showed that the acoustic distance between vocal imitations and referent sounds is not sufficient to account for such performance. They suggested that instead of reproducing the acoustic properties of the referent sound as accurately as vocally possible, vocal imitations focus on a few important features dependent on each particular sound category.

10:45

5aPP11. The perceptual effects of tonewood choice for the back/side plates of the steel-string acoustic guitar. Samuele Carcagno (Dept. of Psych., Lancaster Univ., Lancaster LA1 4YF, United Kingdom, s.carcagno@lancaster.ac.uk), Roger Bucknall (Fylde Guitars, Penrith, United Kingdom), Jim Woodhouse (Dept. of Eng., Univ. of Cambridge, Cambridge, United Kingdom), and Christopher J. Plack (Dept. of Psych., Lancaster Univ., Lancaster, United Kingdom)

The sound qualities of different tonewoods used for the back/side plates of acoustic guitars are the subject of intense debate among guitarists. However, comparisons between different guitars are rarely performed under blinded conditions. We asked a sample of professional, semi-professional, and amateur guitarists (n = 52) to rate the sound quality of six steel-string acoustic guitars, built to be as similar as possible except for the woods used for the back/side plates, which were: Brazilian rosewood, Indian rosewood, mahogany, maple, sapele, and walnut. The guitarists played the guitars in a dimly lit room while wearing welder's goggles to prevent visual identification. The sound quality ratings did not differ significantly between guitars. Test-retest correlations, measured for a subset of guitarists (n = 34), were low, suggesting that the lack of significant rating differences was not due to averaging out contrasting individual preferences. The results of a blinded ABX discrimination test, performed by a subset of guitarists (n = 31), indicate that guitarists could not easily distinguish the guitars by their sound or feel. Overall, the results suggest that the species of wood used for the back/side plates of a steel-string acoustic guitar may play a lesser role in perceived sound quality than is commonly believed.

11:00

5aPP12. Psychoacoustic assessment of auditory impression of the coolness and warmness induced by HVAC noise. Seiji Nakagawa (Ctr. for Frontier Medical Enineering, Chiba Univ., 1-33 Yayoicho, Inage-ku, Chiba, Chiba 263-8522, Japan, s-nakagawa@chiba-u.jp), Takuya Hotehama (Biomedical Res. Inst., National Inst. of Adv. Industrial Sci. and Technol. (AIST), Ikeda, Osaka, Japan), and Masaru Kamiya (DENSO Corp., Kariya, Aichi, Japan)

Noise induced by a heating, ventilation and air conditioning (HVAC) system in a vehicle is an important factor that affects the comfort of the interior of a car cabin. To date, much effort has been devoted to reduce noise levels. However, for example, silence sometimes causes a loss of operation

feeling—we recognize the HVAC doesn't work enough in case the air flow sound is too quiet, even if it works well actually. Therefore, there is a need for a new sound design that addresses the noise problem from a different point of view in addition to the reduction of noise levels. In this study, focusing on the auditory impression of automotive HVAC noise concerning coolness and warmness, listening tests were performed using a paired comparison technique under various conditions of room temperature. Five stimuli were synthesized by stretching the spectral envelopes of recorded automotive HVAC noise to assess the effect of the spectral centroid, and were presented through headphones at 70 dBA. Results show that the spectral centroid significantly affects the auditory impression concerning coolness and warmness; a higher spectral centroid induces a cooler auditory impression regardless of the room temperature.

<div align="center">11:15</div>

5aPP13. On the color of voices: Does good perception of vocal differences relate to better speech intelligibility in cocktail-party settings? Nawal El Boghdady, Deniz Başkent (Dept. of Otorhinolaryngology, Univ. of Groningen, Univ. Medical Ctr. Groningen, Hanzeplein 1, Groningen, Groningen 9700RB, Netherlands, n.el.boghdady@umcg.nl), and Etienne Gaudrain (Auditory Cognition and PsychoAcoust., CNRS UMR5292, Lyon Neurosci. Res. Ctr., Université Lyon 1, Groningen, Netherlands)

Understanding speech in cocktail-party settings poses a challenge for cochlear implant (CI) users. Distortions of vocal cues as they are perceived through the implant have been hypothesized to partially contribute to this difficulty. Vocal cues can be defined along two orthogonal dimensions: the fundamental frequency, F0, related to the pitch, and the vocal tract length (VTL), correlated with speaker size. While deficiency in F0 perception in CI users has been long known, recent research has shown a large deficit in VTL perception. How this contributes to speech intelligibility in the presence of a competing talker (speech-on-speech; SOS) remains unknown. In this study, we investigated the relationship between SOS intelligibility and the sensitivity to F0 and VTL differences (JNDs) in 18 CI listeners using two separate speech tasks. Results indicate a strong correlation between voice cue sensitivity and SOS intelligibility: participants who are more sensitive to small differences in F0 and VTL tend to perform better on the two SOS tasks compared to those who are less sensitive. These results are consistent with the hypothesis that degraded voice cue sensitivity contributes to SOS perception deficits, but other potential explanations, like differences in cognitive processing and attention, cannot yet be ruled out. [**Funding:** The University Medical Center Groningen (UMCG), Advanced Bionics, Netherlands Organization for Scientific Research (NWO).]

<div align="center">11:30</div>

5aPP14. Can a pitch be "sharp," "bright," "large," "narrow," and "high?" Questioning the automaticity of audiovisual correspondences. Laura Getz and Michael Kubovy (Univ. of Virginia, PO Box 400400, Charlottesville, VA 22904, laura.getz@villanova.edu)

Previous research has found that there is an inherent association between auditory and visual dimensions such as the height a pitch and the size of an object. From this, researchers have assumed that such audiovisual correspondences must result solely from bottom-up processing. In a series of studies, we sought to separate bottom-up and top-down effects in the correspondence between pitch and visual size, elevation, spatial frequency, brightness, and sharpness. Using a modified speeded classification task, we asked participants to pair audiovisual dimensions in "compatible" (e.g., high pitch/small circle) and "incompatible" (e.g., high pitch/large circle) conditions. We compared their reaction times across conditions and found that in most cases participants can pair the dimensions in either direction with similar speed and accuracy. We conclude that top-down effects such as task instructions and language knowledge do influence the strength of audiovisual associations. We thus strongly question the assumption of automaticity prevalent in the cross-modal correspondence literature.

<div align="center">11:45</div>

5aPP15. Comparing preferred spectral balances of listeners from four countries. Sungyoung Kim (ECT Eng. Technol., RIT, ENT-2151, 78 Lomb Memorial Dr., Rochester, NY 14623, sungyoungk@gmail.com), Ron Bakker (Commercial Audio, Yamaha Music Europe, Vianen, Netherlands), and Masahiro Ikeda (System Platform Group, Yamaha Corp., Hamamatsu, Shizuoka, Japan)

An influence of cultural background in timbre perception was investigated. The authors have conducted a listening test in the United States, Japan, Korea, and the Netherlands. In the test, each participant was asked to adjust gains of a three-bands equalizer and generate spectral balances of five musical excerpts (four popular songs, one each from four countries and one orchestral music) according to one's preference. All listeners used a same type headphone and audio interface, and repeated the entire process twice. The authors analyzed the consistency of each participant by measuring a total spectral difference (TSD) value between the two trials. Subsequently the data of listeners with less than 45 dB TSD were analyzed through a mixed-design ANOVA. The results showed that the MUSIC factor was not significantly differentiated the reported data ($F(4,86.1) = 2.087$, $p = 0.086$), while a significant interaction existed between GROUP and MUSIC ($F(12,319.3) = 2.58$, $p = 0.004$). In addition, the frequency BAND factor ($F(2,1472) = 28.154$, $p = 0$) and its interaction with GROUP were significant ($F(6, 404.3) = 2.578$, $p = 0.026$). The results implicate that when asked to actively render a preferred timbre balance, the cultural background starts to significantly interact with other variables.

Session 5aSC

Speech Communication: Voice, Tone, Prosody, and Affective and Emotional Qualities of Speech (Poster Session)

Tanya Flores, Chair
Languages and Literature, University of Utah, 255 S Central Campus Dr., LNCO 1400, Salt Lake City, UT 84109

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

Contributed Papers

5aSC1. Effects of perception and vocabulary training of Mandarin tones for native speakers of Japanese: Pre-, post-, and retention test comparison. Shuyi Yang (Graduate School of Intercultural Studies, Kobe Univ., Hyōgo-ken, Kōbe-shi, Nada-ku, Tsurukabuto, 1 Chome−2−1, Kobe 6570815, Japan, syuiyang@gmail.com) and Reiko Akahane-Yamada (ATR, Soraku-gun, Kyoto, Japan)

Native speakers of Japanese were trained to perceive four Mandarin tones, and to semantically and phonetically distinguish Chinese monosyllabic word contrasting in tones. Various tests were administered not only before and after the training period but also forty days after the completion of training. Participants were divided into three groups. First group received perception training first and vocabulary training later. Second group received vocabulary training first and perception training later. The last group, which was the control group, received no training and participated only in testing sessions. The result showed that the accuracy in perception-related-tests improved by perception training, and the accuracy in vocabulary-related-tests improved by vocabulary training. In addition, the effect of training had retained even 40 days after the completion of training. More importantly, the group which received perception training first showed significantly larger improvement from pre-test to post-test than in the group which received vocabulary training first. Note that two groups received the same amount of equivalent trainings in total. Preset results demonstrated the existence of the order effect in foreign language learning. [Work supported by JSPS KAKENHI 23242032.]

5aSC2. What ranges of two-mass model parameters should be used in subject-specific and population-based modeling studies? Douglas Cook (Eng., New York Univ. Abu Dhabi, PO Box 903, New York, NY 10276, prof.laji@gmail.com)

Two-mass models have been used in voice research for over 40 years. It is therefore both surprising and somewhat troubling that there is no firm consensus regarding the values of model parameters that should be used to represent human phonation. A knowledge of the parameter ranges that can (or should) be used is essential for scientifically valid studies involving population-based or subject-specific modeling. In this study, four techniques were used to examine the ranges of two-mass model parameter values that produce behavior representative of human phonation. The first approach involved a review of values that have been used in previous modeling studies. The second approach utilized unrestricted Monte Carlo sampling to examine which ranges can be used to simulate human phonation. The third approach also utilized Monte Carlo sampling, but parameters were restricted based on physical features of the vocal folds. Finally, a reduction or order technique was developed that allows the determination of two-mass model parameters from the physical features of human vocal folds. Finally, results from each of the four methods were compared and contrasted to provide a better understanding of parameter ranges for two-mass models.

5aSC3. A measurement study on voice instabilities during the register transition. Yasufumi Uezu and Tokihiko Kaburagi (Kyushu Univ., 4-9-1, Shiobaru, Minami-ku, Fukuoka-shi, Fukuoka 815-8540, Japan, 3DS14006W@s.kyushu-u.ac.jp)

When one of the dominant harmonics, the fundamental frequency and its harmonic components, is close to the first formant frequency, the effect of the source-filter interaction can induce voice register transition, in which the vocal-fold vibration becomes unstable and the pitch jumps drastically. In this study, we investigated the relationship between the dominant harmonics and the first formant frequency in the modal-falsetto transition to clarify the effect of source-filter interaction. While five subjects performed rising glissandi with /a/ and /i/ vowels, we simultaneously measured their vocal-fold vibration by using electroglottography and the acoustic response from the vocal tract by using the external acoustic excitation method. We analysed temporal patterns of the fundamental and the first formant frequencies in the transition section. We found that the fundamental frequency was regularly in the vicinity of the first formant frequency for the /i/ vowel. Additionally, for the /a/ vowel, it was occasionally observed that the second or third harmonic component was close to the first formant frequency in the transition. Our results indicate that the source-filter interaction is a common factor of the modal-falsetto transition for the participants.

5aSC4. Rhythm segment constitution showing regular periodicity. Shizuka Nakamura (Graduate School of Informatics, Kyoto Univ., 36-1 Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan, shizuka@sap.ist.i.kyoto-u.ac.jp)

To verify the possibility of regular periodicity of English rhythm, each sentence was divided into respective rhythm segments and the properties of its durations were analyzed. Rhythm segment (RhySeg) was defined as a segment including one syllable with a primary/secondary stress to which an adjacent unstressed syllable(s). The following locations of a stressed syllable in RhySeg were compared: forward, semi-forward, middle, semi-back, and back. To reflect the perceptual effect to the RhySeg constitution, the following factors to equally compress all of the unstressed syllables were compared: 0.1-1.0 at an interval of 0.1. To find the RhySeg constitution showing regular periodicity, not only the degree of concentration of the distribution, but the degree of closeness between RhySeg with a secondary stress and 1/2 of that with a primary stress, whose engagement on regular periodicity was indicated in previous studies, was applied as a criterion. Comparative experiments showed the best when the stressed syllable located semi-forward or forward, and the factor was 0.7. Furthermore, the result of harmonic analysis and resynthesis applied to the time function of the average duration of syllables in a sentence indicated periodicity is consisted of the combination of the fundamental and its second harmonic components.

5aSC5. Laryngeal dynamics and vocal fold tissue engineering in a rabbit model. Michael Döllinger (Dept. of Otorhinolaryngology, Head and Neck Surgery, Div. of Phoniatrics and Pediatric Audiol., Univ. Hospital Erlangen, Bohlenplatz, 21, Erlangen, Bavaraia 91054, Germany, michael.doellinger@uk-erlangen.de), Veronika Birk, Stefan Kniesburges (Dept. of Otorhinolaryngology, Head and Neck Surgery, Div. of Phoniatrics and Pediatric Audiol., Univ. Hospital Erlangen, Erlangen, Bavaria, Germany), Christoph Alexiou (Dept. of Otorhinolaryngology, Head and Neck Surgery, Section of Experimental Oncology and Nanomedicine, Univ. Hospital Erlangen, Erlangen, Bavaria, Germany), Olaf Wendler (Dept. of Otorhinolaryngology, Head and Neck Surgery, Experimental ENT Res. Lab. I, Univ. Hospital Erlangen, Erlangen, Germany), Marina Pöttler (Dept. of Otorhinolaryngology, Head and Neck Surgery, Section of Experimental Oncology and Nanomedicine, Univ. Hospital Erlangen, Erlangen, Germany), Anne Schützenberger, and Stefan Dürr (Dept. of Otorhinolaryngology, Head and Neck Surgery, Div. of Phoniatrics and Pediatric Audiol., Univ. Hospital Erlangen, Erlangen, Germany)

Vocal fold surgery, especially due to cancer treatment, yields reduced voice quality and consequently reduced quality of life for patients. Hence, the development of vocal fold implants, to restore missing vocal fold tissue after surgery, is an urgent clinical need. To achieve this, a rabbit model is applied as a first step. Ex-vivo dynamic experiments were performed on twelve rabbit larynges providing normative phonatory data. The larynges were phonated at sustained phonation for different elongation levels at varying subglottal pressures. Laryngeal vibrations, airflow, and acoustics were recorded. Subsequently, for each larynx, a defined area of one vocal fold was resected, simulating the surgical intervention, and were phonated again with the same stimulations. The untreated larynges showed expected behavior regarding flow-pressure relation, acoustics and dynamics. In contrast, the phonatory quality of the resected larynges was significantly reduced showing, as expected, highly disturbed dynamics and acoustics. Parallel, vocal fold fibroblasts were isolated from rabbit larynges and cultured. These cells were treated with superparamagnetic iron oxide nanoparticles enabling their magnetic control. By magnetic tissue engineering three dimensional structures were designed. Next, the applicability of this tissue engineered implant will be tested in the dynamic ex-vivo rabbit model to compare the phonatory outcome.

5aSC6. A study of reliability parameters extracted through voice analysis. Hyung Woo Park (IT, SoongSil Univ., 1212 Hyungham Eng. Bldg. 369 Snagdo-Ro, Dongjak-Gu, Seoul, Seoul 06978, South Korea, pphw@ssu.ac.kr) and Sangmin Lee (Business Administration, SoongSil Univ., Seoul, South Korea)

The human voice is one of the easiest methods for the information transmission between human beings. The characteristics of the voice can be varied by different people and different situations, such as utterance speed, pitch tone, vocal organ features, and the gender. Moreover, the voice can be used as a factor for deciding the personal credit rating scores. The reliable parameters of speech signal can be extracted from different characteristics of the spoken information. In this paper, we collected the voices from people who has relatively high personal credit scores, and analyzed them.

5aSC7. Acoustic similarities among female voices. Patricia Keating and Jody Kreiman (Dept. of Linguist, UCLA, Los Angeles, CA 90095-1543, keating@humnet.ucla.edu)

Little is known about how to characterize normal variability in voice quality within and across utterances from normal speakers. Given a standard set of acoustic measures of voice, how similar are samples of 50 women's voices? Fifty women, all native speakers of English, read 5 sentences twice on 3 days—30 sentences per speaker. The VoiceSauce analysis program estimated many acoustic parameters for the vowels and approximant consonants in each sentence, including F0, harmonic amplitude differences, harmonic-to-noise ratios, formant frequencies. Each sentence was then characterized by the mean and standard deviation of each measure. Linear discriminant analysis tested how well each speaker's set of 30 sentences could be acoustically distinguished from all other speakers' sentences. Initial work testing just 3 speakers from this sample found that the speakers could be completely discriminated (classified) by these measures, and

largely discriminated by just 2 of them. Such a simple result is not expected for the larger sample of speakers. We will present results concerning how successfully speakers can be discriminated, how well different numbers of discriminant functions do, and which acoustic measures do the most work. Implications for recognition by listening will be discussed. [Work supported by NSF and NIH.]

5aSC8. A study on prediction of end-of-utterance by prosodic features and phrase-dependency structure in spontaneous speech. Yuichi Ishimoto (Ctr. for Res. Resources, National Inst. for Japanese Lang. and Linguist, 10-2 Midoricho, Tachikawa, Tokyo 190-8561, Japan, yishi@ninjal.ac.jp), Takehiro Teraoka, and Mika Enomoto (School of Media Sci., Tokyo Univ. of Technol., Hachioji, Tokyo, Japan)

This study is aimed at predicting the end of utterance by prosodic features and syntactic structure for spontaneous speech. In spontaneous everyday conversation, participants must predict the ends of utterances of a speaker to perform smooth turn-taking. We consider that they utilize not only syntactic factors but also prosodic factors for the end-of-utterance prediction because of the difficulty of prediction of a syntactic completion point in spontaneous Japanese speech. In previous studies, it was observed that prosodic factors changed such that the general fundamental frequency of utterance declined gradually toward the end of an utterance, and the intensity decreased significantly in the final accentual phrase. However, it is not clear what prosodic features support the prediction. We focused on dependency structure among bunsetsu-phrases as the syntactic factor and investigated the relation between the phrase-dependency and prosodic features based on a spontaneous Japanese conversation corpus. The results showed that the average fundamental frequency and the average intensity for accentual phrases did not decline until the modified phrase appeared. This suggests the possibility that prosodic changes and phrase-dependency relations inform the hearer that the utterance is approaching its end.

5aSC9. Working memory facilitates the detection and correction of feedback errors in vocal pitch regulation. Hanjun Liu, Zhiqiang Guo, and Xiuqin Wu (Rehabilitation Medicine, The First Affiliated Hospital of Sun Yat-sen Univ., 58 Zhongshan 2nd Rd., Guangzhou, Guangdong 510080, China, lhanjun@mail.sysu.edu.cn)

In speech processing, information related to the speech motor command and sensory re-afference can be stored in working memory (WM) within a feedback circuit and recalled when needed to adjust the motor activity. Whether WM facilitates the online monitoring of speech motor control, however, remains unclear. The present event-related potential study sought to examine the impact of WM on the auditory-motor processing of pitch feedback errors. Participants sustained a vowel phonation while hearing their voice pitch-shifted +200 or +500 cents five times. In the WM task, participants were asked to determine whether the sequence of 5 pitch shifts was different or not between two consecutive vocalizations. In the control task, they did nothing but maintaining their vocalizations steady when exposed to pitch-shifted auditory feedback. The behavioral results revealed a significant increase of vocal responses in the WM task as compared to the control task. At the cortical level, the WM task elicited significantly larger N1 responses and smaller P2 responses than the control task. Taken together, these findings demonstrate the influence of WM on the neurobehavioral responses to pitch-shifted voice auditory feedback, indicating that WM can facilitate the detection and correction of pitch feedback errors in vocal motor control.

5aSC10. Effect of level difference between left and right vocal folds on phonation: Physical experiment and theoretical study. Ryo Shimamura (Graduate School of Sci. and Eng., Ritsumeikan Univ., Noji-higashi, 1-1-1, Kusatsu, Siga 525-0058, Japan, rt0021hf@ed.ritsumei.ac.jp) and Isao T. Tokuda (Graduate School of Sci. and Eng., Ritsumeikan Univ., Kusatsu, Shiga, Japan)

The vocal folds, which are constituted by muscles covered with a mucous membrane, generate a primary sound called the voice source, as airflow passes them. In some voice disorders, asymmetry between left and right vocal folds was observed. We focus on level difference, which is defined as the distance between the upper surfaces of the bilateral vocal

folds in the inferior-superior direction and is caused by symptom of such disorders. Physical models of the vocal folds were utilized to study the effect of the level difference on the phonation threshold pressure. For three types of different self-oscillating synthetic models, our experiments reveal that the phonation threshold pressure increases significantly as the level difference is extended. Furthermore, based upon a small amplitude approximation of the vocal fold oscillations, a theoretical formula was derived for the phonation threshold pressure. Our theory was in good agreement with the experiments, especially when the phase difference between the left and right vocal folds is not too large. From these results, we conclude that the level difference affects voice production, therefor the effect of the vocal fold geometry needs to be taken into account for the observation of voice disorders.

5aSC11. Psychoacoustic roughness as creaky voice predictor. Julián Villegas (Univ. of Aizu, The University of Aizu, Tsuruga, ikki-machi, Aizu Wakamtsu, Fukushima 965-8580, Japan, julian@u-aizu.ac.jp), Jeremy Perkins (Univ. of Aizu, Aizu-Wakamatsu, Japan), and Seunghun J. Lee (Int. Christian Univ., Tokyo, Japan)

The use of psychoacoustic roughness as a predictor of creaky voice is reported. Roughness, a prothetic sensation elicited by rapid changes in the temporal envelop of a sound (15-300 Hz), shares qualitative similarities with a kind of phonation known as vocal fry or creakiness. When a creakiness classification made by trained linguists was used as a reference, a classifier based on an objective temporal roughness model yielded results similar to an artificial neural network-based predictor of creakiness, but the former classifier tended to produce more type I errors. We also compare the results of the roughness-based prediction with those predicted by samples of three populations who use creakiness contrastively in different degrees: Japanese (where creakiness is not systematically used for phonetic contrast), Mandarin (where creakiness is used as a secondary cue), and Vietnamese (where creakiness is used as a phonetic contrast between tones). The roughness-based classification seems to better agree with classifications made by the untrained listeners. Our findings suggest that extreme roughness values (>4 asper) in combination with local prominences on the roughness temporal profile of vocalic segments could be used for classification of creaky intervals in running speech.

5aSC12. A study on the intonation of Japanese modal particles "ne" and "yo". Haixia Zhang and Ayako Shirose (Graduate School of Education, Tokyo Gakugei Univ., 4-1-1,Nukuikita-machi, Koganei, Tokyo 184-8501, Japan, wsbxk544@yahoo.co.jp)

The Japanese sentence-final particles "ne" and "yo" represent modality. These particles are assumed to have several meanings, but the definition of the meanings of them has been an open question. In particular, the relationship between the meaning of these particles and their intonation contours remains unresolved. Therefore, to clarify this, we investigated the meaning of the particles "ne" and "yo" in actual utterances using a questionnaire and analyzed the pitch contours of the utterances acoustically. A Standard Japanese speaker was asked to produce 36 utterances with "ne" or "yo" particles controlling three types of pitch contours, falling, rising and flat. First, we asked 44 Japanese university students (aged 18-24) to judge the "naturalness" of the obtained utterances and to describe their meaning. The results revealed that the particle "ne" tended to have a negative connotation in the case of the falling intonation. Second, the pitch contours of final two syllables of the utterances were analyzed acoustically. The results indicated that the intonation contours differed between the two particles; the rising range was wider in "ne" than "yo". This was assumed to reflect the different meaning between the two particles. Speech corpus data will be examined to verify above results.

5aSC13. Mandarin tone recognition based on unsupervised feature learning from spectrograms. Charles Chen, Razvan Bunescu (Elec. Eng. and Comput. Sci., Ohio Univ., 1 Ohio University, Athens, OH 45701, lc971015@ohio.edu), Li Xu (Commun. Sci. and Disord., Ohio Univ., Athens, OH), and Chang Liu (Elec. Eng. and Comput. Sci., Ohio Univ., Athens, OH)

In tone languages, such as Mandarin Chinese, a syllable with different tones conveys different meanings. Previously, Mandarin tone recognition based on Mel-frequency cepstral coefficients (MFCCs) and Convolutional Neural Networks (CNN) was examined and the results outperformed the model of conventional neural network using manually edited F0 data. In the present study, Mandarin tone recognition based on spectrograms, instead of MFCCs, was explored. Unsupervised feature learning was applied to the unlabeled spectrograms directly with a denoising autoencoder (dAE). Then, the model convolved the labeled spectrograms with the learnt "sound features" and produced a set of feature maps. A dataset that consisted of 4500 monosyllabic words collected from 125 children was used to evaluate the recognition performance. Compared with methods based on MFCCs, there are more parameters to train in the new approach based on spectrograms. As a result, the new model might better capture the statistical distribution in the original data. Therefore, the new approach, with unsupervised feature learning, could perform even better than previous methods based on MFCCs or those based on the extracted F0 information. The advantages and shortcomings of various approaches for lexical tone recognition will be discussed. [Work supported in part by the NIH NIDCD Grant No. R15-DC014587.]

5aSC14. Distribution of accentual phrase medial tones in Seoul Korean. Hyun Ji Yoo and Sun-Ah Jun (Linguist, UCLA, University of California, Los Angeles, Los Angeles, CA 90095, yhj8290@ucla.edu)

According to the model of Korean (Seoul) intonation (Jun 1993, 2007), the underlying tonal pattern of an Accentual Phrase (AP) is either HHLH (if the AP-initial segment is aspirated or tense) or LHLH (elsewhere). The second tone, H (= +H in K-ToBI; Jun 2000), or the third tone, L (= L+ in K-ToBI), or both are known to be often undershot when the AP is shorter than 4 syllables, but so far factors determining the distribution of these medial tones have not been investigated. This study aims to find predictors of AP-medial tone realization based on K-ToBI labeled speech of various speech styles. Preliminary results show that in general L+ is more frequently undershot than +H regardless of the AP size. This pattern was clearer in L-initial APs than H-initial APs. Furthermore, the type of the AP-medial tone (when IP-final) showed an interaction with the type of immediately following Intonational Phrase boundary tones: L+ was more often undershot than +H before L%, but the opposite was true before H(L)%. The results suggest that the tonal context adjacent to the AP-medial tone plays an important role in the distribution of AP-medial tones. More predictors will be discussed at the meeting.

5aSC15. On the necessity of using 3D imaging of vocal fold vibration. David Berry (Surgery, UCLA, 31-24 Rehab, Los Angeles, CA 90095-1794, daberry@ucla.edu)

Physical mechanisms of regular and irregular vocal fold vibration were first studied using a computational model of vocal fold vibration. Later physical mechanisms of vocal fold vibration were studied in laboratory hemilarynx studies, where the medial surface of the vocal folds were imaged. While the method was also extended to clinical studies of vocal vibration, the method exhibited significantly less interpretive power in the clinical applications in which the superior surface of the vocal folds was imaged in 2D. Our hypothesis is that the interpretive power of the method of empirical eigenfunctions in studying the superior surface dynamics of the vocal folds will be significantly increased if one performs 3D imaging of the dynamics versus the standard 2D imaging. To test this hypothesis, the method of empirical eigenfunctions was employed on the same finite element model used in previous computational experiments. Our results confirm the hypothesis that the interpretive power of the method of empirical eigenfunctions was significantly increased in studying the superior surface dynamics of the vocal folds when 3D imaging was employed.

3394   J. Acoust. Soc. Am., Vol. 140, No. 4, Pt. 2, October 2016

5th Joint Meeting ASA/ASJ   3394

5aSC16. Relations on prosody of Japanese back-channeling word "hai" and listener's impression: An investigation with synthesized voices. Tetsu Tanahashi and Tetsuro Kitahara (Graduate School of Integrated Basic Sci., Nihon Univ., 3-25-40, Sakurajosui Setagaya-ku, Tokyo 156-8550, Japan, tanahashi@kthrlab.jp)

The back-channeling word "hai" in Japanese is basically used in positive contexts, but it may give a negative impression to the listener if an inappropriate prosody is used. The same thing will occur in human-system dialogues using synthesized voices. Therefore, we investigated the relations on the prosody of synthesized voices of "hai" and the listener's impression. An experiment was conducted using the semantic differential method. Synthesized voices of "hai" were presented with different speaking rates (slow and fast) and different fundamental frequencies (low and high). 16 participants (university students, 8 males and 8 females) rated them on 28 adjective scales. The results were analyzed with the factor analysis, so that four factors were extracted. For the first factor, the loadings of "definite", "calm", "active",etc., were higher than 0.5, so the first factor is considered to represent positiveness. For the second factor, the loadings of "definite," "violent," "cold," and etc., were higher than 0.5, so the second factor is considered to represent unfriendliness. The first factor tends to be higher as the fundamental frequency is lower. The second factor tends to be lower as the speaking rate is higher.

5aSC17. The nature of tone sandhi in Sixian Hakka. Jie Zhang (Linguist, The Univ. of Kansas, 1541 Lilac Ln., Blake Hall, Rm. 427, Lawrence, KS 66045, zhang@ku.edu), Hanbo Yan (Shanghai Int. Studies Univ., Shanghai, China), Yuwen Lai, and Shao-ren Lyu (National Chiao-Tung Univ., Hsinchu, Taiwan)

Sixian Hakka is a Hakka dialect spoken in Taiwan. The language has four contrastive tones on non-checked syllables: 24, 11, 31, and 55. But before 24 or 55, a tone sandhi pattern changes 24 to 11, neutralizing the tonal contrast to three. We report two experiments that tap into the nature of this tone sandhi pattern in this paper. The first is a nonce-probe test, in which participants produced novel words that met the sandhi environment (/24-24/, /24-55/), and the f0 of the first syllable was compared to the f0 produced in real words to test the productivity of the sandhi. The second is an auditory lexical decision experiment with auditory priming, which tested whether disyllabic tone sandhi words were primed by monosyllables that shared the base tone or sandhi tone of the first syllable. Both experiments were conduced in Miaoli, Taiwan. We will report data from 15 participants for the first experiment and 32 for the second experiments. Together with experimental studies on various types of tone sandhi patterns elsewhere (e.g., Mandarin, Taiwanese), the results here will shed further light on how speakers internalize complex tonal alternation. The study also provides experimental data on an understudied and disenfranchised language.

5aSC18. A computational study of vocal fold dehydration during phonation. Liang Wu and Zhaoyan Zhang (Dept. of Head and Neck Surgery, Univ. of California, Los Angeles, 1000 Veteran Ave, 31-24 Rehab Ctr., Los Angeles, CA 90095, zyzhang@ucla.edu)

Vocal fold dehydration during phonation is investigated in a continuum model of the vocal folds. Based on the linear poroelastic theory, the model simulates water movement inside the vocal folds during phonation, water exchange between the vocal folds and the surface mucosal layer through the epithelium layer, surface water accumulation and loss to the glottal airway, and water resupply from blood through the lateral boundary. Parametric studies are conducted to investigate water loss within the vocal folds after 5 minutes of phonation at different voice conditions. The results show that with normal water resupply from the blood, water loss within the vocal folds increases with greater vibration amplitude and higher epithelium permeability. At very large vibration amplitudes, the water loss within the vocal folds can be as high as 3%, which may severely affect body functions and possibly phonation characteristics and vocal effort. Reduced water resupply from blood further increases the degree of dehydration. In contrast, water loss in the surface mucosal layer is an order of magnitude higher than water loss within the vocal folds, indicating the surface dehydration level is likely not a good indicator of the degree of dehydration within the vocal folds.

5aSC19. Where is the beat? Lauryn Zipse (Inst. for Health Professions-MGH, Boston, MA), Olivia Murton, Katie Dunn, and Stefanie Shattuck-Hufnagel (Speech Commun. Group, MIT, 36-511 MIT, 77 Mass Ave., Cambridge, MA 02139, sshuf@mit.edu)

The ability to find the beat in a sequence of auditory events may be linked to the ability to learn vocal communication, raising the question of how beat structure in speech events relates to that in other event sequences. We conducted a series of entrainment experiments designed to compare spoken syllable repetition with tapping. Producing taps to a periodic string of auditorily-presented spoken /pa/ syllables resulted in the tap falling between the release burst for the /p/ and the onset of voicing for the vowel. This is consistent with participants intending to align their taps with the vowel onsets but exhibiting the well-documented Negative Mean Asynchrony (NMA) effect, such that the taps precede their "target." The finding of alignment with the voicing onset is reminiscent of a large body of work on the P-center in repeated spoken syllables. In contrast, producing repeated utterances of the syllable /pa/ to an auditorily-presented click train resulted in *coincident occurrence of the release burst with the click,* putting the voicing onset after the stimulus (inconsistent with a voicing-onset target with associated NMA). These findings indicate that the location of the beat may differ across tasks involving speech perception and various types of speech production.

5aSC20. An empirical model linking intrinsic laryngeal muscle activation to vocal fold geometry and stiffness. Zhaoyan Zhang, Dinesh K. Chhetri, and Andrew M. Vahabzadeh-Hagh (Dept. of Head and Neck Surgery, Univ. of California, Los Angeles, 1000 Veteran Ave, 31-24 Rehab Ctr., Los Angeles, CA 90095, zyzhang@ucla.edu)

Although vocal fold vibration and the resulting voice production depend on the geometry and mechanical properties of the vocal folds, control of voice is achieved through laryngeal muscle activation, which postures the vocal folds into desired geometry and mechanical properties. Such posturing process is known to cause simultaneous changes in vocal fold position, geometry, stiffness, and tension, which has been hypothesized to be a possible origin of vocal registers. Understanding such co-variations in vocal fold geometry and mechanical properties is thus critical to our understanding of human voice production and control, and the development of muscularly controlled, computationally efficient models of phonation. In this study, an empirical model is developed to link the glottal gap, vocal fold length, vocal fold vertical thickness, and the mechanical properties of the vocal fold to the activation levels of the cricothyroid, thyroarytenoid, lateral cricoarytenoid, interarytenoid, and posterior cricoarytenoid muscles, based on data from *in vivo* larynx experiments. Combined with a continuum model of phonation, this model is used to investigate muscular control of pitch and vocal registers and the preliminary results are discussed.

5aSC21. Empirical relationship in excised canine larynx. Sid M. Khosla, Liran Oren (Otolaryngol., Univ. of Cincinnati Academic Health Ctr., 231 Albert Sabin Way, 0528, Cincinnati, OH 45208, khoslasm@uc.edu), and ephraim gutrmark (Aerosp. Eng., Univ. of Cincinnati, Cincinnati, OH)

Our lab uses particle imaging velocimetry to measure intraglottal pressure and geometry. Although we have described relationships between maximum divergence angle (MDA) and maximum negative pressure during closin (MNP) for individual experiments, we have not yet looked at the combined data for over 15 larynges without a vocal tract. In this presentation, plots will be shown for the following relationships: MDA vs MNP, MNP vs vortex strength, MNP vs maximum closing speed and maximum vocal fold closing speed vertical stiffness gradient vs divergent angle. Correlation coefficients and generalized empirical relationships will be shown. This will allow approximations for finding variable that are hard to calculate computationally. For example, it is much easier to calculate divergence angle than vortex strength. Comparison of empirical and computational relationships will be described.

5a FRI. AM

5aSC22. Tone-intonation interaction in context in Thai. Amber B. Camp (Linguist, Univ. of Hawaii at Manoa, 1890 East West Rd., Moore Hall #569, Honolulu, HI 96822, acamp@hawaii.edu)

The acoustic realization of lexical tone is influenced by sentence-level intonation. For example, F0 measurements show that a phonologically falling tone in Thai is warped when overlaid with a falling intonational contour, causing F0 to fall further and to begin its decline at an earlier timepoint than when influenced by other contours. This categorical perception study, which includes both identification and discrimination tasks, uses a nine-step continuum of target words created with naturally produced lexical tone endpoints, presented within two different naturally produced sentence frames. Specifically, these tasks investigate the perception of high and falling tones in sentence-medial and sentence-final positions in Thai. This test of stimulus continua within natural sentence contexts reveals the listeners' context-mediated perception of deviation from a canonical lexical tone, shedding light on the interaction and relative weights of both tone and intonation perception in natural speech, and also offering insight into the mechanism with which language users process suprasegmental information.

5aSC23. Acoustic properties of stress in Kubeo. Matthew K. Gordon (Dept. of Linguist, UC Santa Barbara, UC Santa Barbara, Santa Barbara, CA 93106, mgordon@linguistics.ucsb.edu), Thiago C. Chacon (Univ. of Brasilia, Brasilia, Brazil), Kaveh Varjoy (Linguist, UC Santa Barbara, Santa Barbara, CA), and Jullie A. Ferreira (Univ. of Brasilia, Brasilia, Brazil)

This paper explores the realization of stress in Kubeo, a Tukanoan language from the Northwest Amazonia with both stress and tone (H vs. HL). The first H tone of a word docks on the primary stressed syllable, which by default is the second syllable except in words with lexical initial stress. Secondary stress occurs every two syllables to the right of the primary stress. F0, duration, intensity and the first two formants were measured for vowels in a corpus of 150 words ranging from one to six syllables. Results from two speakers indicate that primary stress is associated with the greatest duration and most peripheral vowel qualities. Unstressed pre-tonic syllables are phonetically more prominent than their post-tonic counterparts along all dimensions except F0, which depends on lexical tone. In post-tonic unstressed syllables, duration, intensity, and vowel quality are all dependent on lexical tone. Secondary stress is less salient due to interference from lexical tones but appears to be associated with longer duration and more peripheral vowels relative to unstressed syllables. Our results enrich typological knowledge of both the acoustic realization of stress in tone languages and the interplay of tonal and metrical phonology in languages with complex word level prosody.

5aSC24. An intonational investigation of five sentence types in Kannada. Kirsten Helgeson (Linguist, Univ. of Hawaii at Mānoa, 1890 East-West Rd., 569 Moore, Honolulu, HI 96822, kjhelg@hawaii.edu)

Kannada (ISO 639-3 kan) is a Dravidian language spoken by almost 40 million people, primarily in Karnataka state in southern India. It is the official language of Karnataka and one of the scheduled languages of India. Despite its prominence, there is little published in English on the language's intonation systems. This project furthers knowledge of Kannada intonation that is available to speakers of English, with implications for pedagogy and cross-cultural communication. Two female and two male native speakers were recorded in dyads, participating in a scripted conversation containing five types of sentences. F0 contours were examined to determine intonational tunes for different sentence types.

5aSC25. Cue weighting in the tonal register contrast of Jiashan Wu. Bing'er Jiang and Meghan Clayards (McGill Univ., 1085 Dr. Penfield, Montreal, QC H3A 1A7, Canada, binger.jiang@mail.mcgill.ca)

Chinese Wu dialects are known to have two tonal registers, where the lower register is realized with lower pitch and breathy phonation and the upper register is realized with higher pitch and modal phonation. In Jiashan Wu, the falling tone is realized differently in the two registers: the pitch contour of the upper register is slightly steeper than the lower register. This study investigates how speakers of Jiashan Wu weight the three cues (i.e., breathiness, pitch height, pitch contour) in the register contrast. We recorded

two words /ka/ from the upper and lower register and created stimuli varying in both dimensions (5 steps pitch height x 5 step breathiness = 25 stimuli) and imposed the two contours on all stimuli. 28 native listeners performed a forced-choice categorization task on 5 repetitions of each stimuli in random order. A mixed effect logistic model shows that all three factors affect categorization, and that pitch contour is the most important cue and breathiness the least. Moreover, the effect of breathiness was smaller with higher pitches and a steep contour, and the effect of pitch height is smaller with a steep contour. Data are being collected comparing Jiashan and Shanghai dialects.

5aSC26. Acoustic correlates of prominence in Besemah (Malayic, Indonesia). Bradley McDonnell (Dept. of Linguist, University of Hawai'i at Mānoa, 1890 East-West Rd., 569 Moore, Honolulu, HI 96822, bradley.mcdonnell@gmail.com)

This paper examines the acoustic realization of word-level stress and phrase-level prominence (i.e., pitch-accents, boundary tones) in Besemah, a little-described Malayic language of southwest Sumatra, Indonesia. There has been much disagreement over the status of word-level stress in the languages of western Indonesia, particularly with regards to well-known varieties of Malay (i.e., Standard Malay-Indonesian). Utilizing acoustic cues, word-level stress has been claimed to be present by some (e.g., Adisasmito-Smith & Cohn 1996), though studies involving perception experiments have questioned this position (e.g., van Zanten & van Heuven 2004). All of these studies, however, are complicated by significant influence from substrate languages like Javanese, which apparently lack word-level stress (Goedemans & van Zanten 2007). As a follow up to McDonnell (forthcoming), which found that word-level stress in Besemah falls on the final syllable of the word and is cued by increased intensity, the present study reports the results of a more complex sentence completion task where six speakers of Besemah (3 male, 3 female) uttered target words in different frames that vary along two dimensions: [±focus] and [±final position in utterance]. The study aims to show the acoustic correlates of prominence in Besemah, teasing apart word-level stress, pitch-accent, and boundary tones.

5aSC27. Modeling the cross-linguistic variations of tonal systems. Jianjing Kuang (Linguist, UPenn, 255 S 36th St., Philadelphia, PA 19104, kuangj@sas.upenn.edu)

This study aims to simulate the cross-linguistic variations of tonal systems with low dimensional models. Individual syllables of Mandarin, Yoruba, southern Yi, and Hmong were retrieved from existing speech corpora. Voice quality measures as well as f0 were extracted for all data. For each language, Functional Principal Component Analysis (FPCA) is used to parameterize the f0 contours of individual syllables. It is found that three FPCs are sufficient to account for more than 90% of the variance for pitch contours for all languages. Moreover, the same FPCs are shared among tonal languages and appear to be phonologically meaningful: FPC1 is related to the pitch range; FPC2 is related to the direction of the contour (e.g. rising or falling); FPC3 is related to more complex contours such as dipping and convex. Tone classification models were built for each language, and both voice quality measures and f0 PCs were fitted into the models. Voice quality cues significantly improve the accuracy of the predictions, but different languages vary in the relative importance of voice quality cues. This study provides a new effective way to computationally model the cross-linguistic variation of tonal systems, and has practical implications for speech synthesis.

5aSC28. Distinguishing ambiguity: Cantonese speakers using prosody to disambiguate surface identity in syntax. Elaine Lau (Linguist, Univ. of Hawai'i at Manoa, 7E, Block 3, Site 4, Whampoa Garden, Kowloon 00000, Hong Kong, elau@hawaii.edu)

This paper explores how prosodic and syntactic structures correspond in Cantonese. We consider how native speakers disambiguate object relative clauses (ORC, 1a) and main clauses (MC, 1b), which have a surface identity due to the homonymy between the relative clause marker (go2+CL) and the demonstrative go2+CL 'that'. (1) a. [RC naam4zai2 sek3gan2 ₋] go2 go3 neoi5zai2 boy kiss-PROG that CL girl "the girl that the boy is kissing" b.

naam4zai2 sek3gan2 go2 go3 neoi5zai2 boy kiss-PROG that CL girl "The boy is kissing that girl." Given fundamental frequency (F_0) is used to mark lexical contrasts in the language, it is interesting what other prosodic cues may be used for different parses of the identical string of constituents. Six speakers are recruited for an elicited production study. Pictures designed to elicit ORC vs. MC were presented in random order. Results of acoustic analyses revealed that Cantonese speakers mainly use duration, rather than F_0, to disambiguate the two constructions, reminiscent of focus prosody where F_0 is found to play a limited role. The theoretical implications of these results in relation to the interplay between syntax and prosody are discussed.

5aSC29. Production of neutral tones in three Mandarin dialects. Yang Li (Univ. of Cambridge, Cambridge, United Kingdom) and Arthur L. Thompson (Dept. of Linguist, The Univ. of Hong Kong, Hong Kong, Hong Kong, arthurlewisthompson@gmail.com)

Neutral tone (NT) in Mandarin refers to syllables with shorter duration and whose pitch realisations depend on preceding syllables (Chao 1968; Lee and Zee 2008). There are few acoustic studies on NT in Mandarin dialects outside the standard, and on how NT interacts with sentence modality and prosodic position. We report results from a reading task with 3 speakers each from Chongqing, Kunming, and Nanjing (N = 9). A set of 8 syllables with potential NT and control lexical tones are placed in phrase-medial and final positions and in statements and echo questions. Pitch contours of entire utterances are analysed to determine the variation in pitch height, slope and tonal coarticulatory influence on target syllables. We find NT-like behaviour in Chongqing and Kunming, despite the longstanding claim of their absence (Luo and Wang 1981; Zhai 1996): e.g., in Chongqing, NT syllables are high-falling medially, but undergo more pitch reduction than control and exhibit obligatory tonal spreading from preceding tone regardless of speech rate. Nanjing NT interacts strongly with modality: statement-final NT is almost invariably realised as strong creak if preceding tone is mid-falling. We also report hitherto unreported variation and hypothesise on the impact of standard Mandarin on NT in these dialects.

5aSC30. Two types of intermediate phrase boundary tones in Korean. Hyunah Ahn (Univ. of Hawaii at Manoa, 1890 East-West Rd., #570, Honolulu, HI 96822, hyunah.ahn@hawaii.edu)

The current study shows that there are two types of Intermediate Phrase (ip) boundary tones, distinctively used for focus marking and syntactic phrasing in spoken Korean. Jun (2011) argued that focused ip (ipF) are characterized by raised initial high (+H) tone and dephrasing of subsequent materials and syntactic ip(ipS) by a phrasal juncture (larger than AP but small than IP) on the phrase-final boundary and partial dephrasing of the following phrase. To empirically test the claim, an experiment was conducted where participants were given an ambiguous sentence to compose a story with. The ambiguous construction was made up of an overt NP, a VP and a covert NP. The ambiguity depended on where the covert NP goes. If the word order is a. "overt NP + covert NP + VP," the syntactic juncture between the overt NP and the VP is larger than when the covert NP is placed before the overt NP (b. covert NP + overt NP + VP). In stories where the syntax of the sentence was a, the likelihood of ipS was significantly larger than when the syntax was b (p <.05). The distribution of ipF was related to the information structure of the story (operationalized as the location of the sentence within the story) (p <.01). The results show that the two different types of intermediate phrase (ip) boundary tones are used in real-time production of spoken Korean. The findings provide empirical evidence for Jun's (2011) claim for the Korean prosodic hierarchy.

5aSC31. Vocal fry in realistic speech: Acoustic characteristics and perceptions of vocal fry in spontaneously produced and read speech. Sarah T. Irons and Jessica E. Alexander (Psych., Centenary College of Louisiana, 2911 Centenary Blvd., Shreveport, LA 71104, jalexander@centenary.edu)

There has been a great deal of recent interest in vocal fry, both in production and perception. However, much of the scientific literature that has used naturally produced fry has focused on speech elicited through reading, rather than spontaneous speech. The current study compares reading with spontaneous speech elicited in various ways for both male and female speakers, recorded in dyads. The speakers were asked to teach their partner information, give instructions, and describe their qualifications for a scholarship, with reading controls for each task. Surprisingly, we found more vocal fry (in both proportion of words and words/duration) for men than women. Men differed in the amount of fry across task types. Interactions of speaker sex, task type, and fry were examined for measures of F0, pitch range, intensity level, jitter, shimmer, and HNR. Listener perceptions of speech with and without vocal fry were also obtained. Task type affected listener perceptions of speech with vocal fry. Vocal fry in spontaneously produced speech seems to differ from fry in speech produced during reading, in quantity, acoustics, and listener perceptions. Additionally, acoustic measures suggest that vocal fry may be intimately tied to decreased vocal effort across task types and speaker sex.

5aSC32. A longitudinal study of within-individual variation in human voice pitch. Katarzyna Pisanski (Mammal Vocal Commun. & Cognition Res. Group, Psych., Univ. of Sussex, School of Psych., Pevensey Bldg. Rm. 2A13, Falmer, Brighton BN1 9QH, United Kingdom, K.Pisanski@sussex.ac.uk), Meddy Fouquet (Mammal Vocal Commun. & Cognition Res. Group, Psych., Univ. of Sussex, Saint-Etienne, France), Nicolas Mathevon (Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI CNRS UMR 9197, Univ. of Lyon/Saint-Etienne, Saint-Etienne, France), and David Reby (Mammal Vocal Commun. & Cognition Res. Group, Psych., Univ. of Sussex, Brighton, United Kingdom)

Individual differences in human voice pitch (fundamental frequency, $F0$) have evolutionary relevance. Fundamental frequency indicates the sex, age, and even dominance of the speaker, and influences a host of social assessments including mate preferences. Yet, due to the almost exclusive utilization of cross-sectional designs in previous work, it remains unknown whether individual differences in $F0$ emerge before or after sexual maturation, and whether $F0$ remains stable throughout a person's lifetime. In our study, we tracked within-individual variation in the $F0$s of male and female speakers whose voices were recorded from childhood into adulthood. Voice recordings were extracted from digital archives. Our results corroborate those of earlier cross-sectional studies indicating a sharp decrease in male $F0$ at puberty resulting in the emergence of sexual dimorphism in adult $F0$. Critically, our results further revealed that men's pre-pubertal $F0$ strongly predicted their $F0$ at every subsequent adult age, and that $F0$ remained remarkably stable within-individuals throughout their adulthood. These findings suggest that adult men's voice pitch may be linked to pre-natal/pre-pubertal androgen exposure and may function as a reliable and stable signal of mate quality, with implications for our understanding of the developmental mechanisms, adaptive functions, and social perception of human voice pitch.

5aSC33. Surface categorical variation In the prosodic marking of English focus. Adam J. Royer and Sun-Ah Jun (Linguist, UCLA, 3125 Campbell Hall, Los Angeles, CA 90095-1543, ajroyer@ucla.edu)

This study investigates the variability of focus prosody in English. In intonational phonology of English, focus is known to be marked with a L + H* pitch accent on a focused word followed by deaccenting (Ø) post-focus words (Beckman & Pierrehumbert 1986). 80 speakers participated in a picture naming task where they named images (e.g., blue ball, yellow chair) they saw using the phrase "There was a *Adjective1 Noun1*, but now there's a *Adjective2 Noun2*." Contrastive focus (CF) trials elicited narrow focus on the *Adjective2* and Non-contrastive (NC) ones did not. For "a *Adjective2 Noun2*", f0 contours were ToBI annotated and various acoustic measurements were taken. Preliminary results indicate that in the CF condition, *Adjective2* was most often marked by L + H*, but it is used only 40% of the time. In the other cases, *Adjective2* carried either H*, H + !H*, L + !H*, or !H*. The tune used on *Adjective2 Noun2* was mainly L + H* Ø, but L + H* !H*, H* Ø, and H + !H* Ø also occurred. These results raise questions about previous claims on phonological marking of focus and the level of tonal representation: what is the phonological status of the various pitch accent types used to mark focus? Are they distinctive or allotones of L + H*?

5aSC34. Aeroacoustic sound generation during phonation in a synthetic larynx model. Stefan Kniesburges (Dept. of Otorhinolaryngology, Head and Neck Surgery, Div. of Phoniatrics and Pediatric Audiol., Univerity Hospital Erlangen, Cauerstrasse 4, Erlangen, Bavaria 91058, Germany, stefan.kniesburges@uk-erlangen.de), Alexander Lodermeyer, Matthias Tautz, Stefan Becker (Dept. of Process Machinery and Systems Eng., Friedrich-Alexander Univ. Erlangen-Nürnberg, Erlangen, Germany), and Michael Döllinger (Dept. of Otorhinolaryngology, Head and Neck Surgery, Div. of Phoniatrics and Pediatric Audiol., Univerity Hospital Erlangen, Erlangen, Germany)

The fluid flow within the human larynx plays an essential role in the fluid-structure-acoustic interaction during voice production. This study addresses the supraglottal flow field downstream of aerodynamically driven, synthetic vocal folds based on the M5 model. The larynx replica is designed to provide full optical access to the flow region. Two different approaches based on Particle-Image-Velocimetry (PIV) were applied for measuring the flow: Phase-averaged (PA) and High-speed (HS) PIV. Beside a comparison of the supraglottal flow field, the acoustic sources were calculated based on both PIV approaches. Furthermore, the simulated far-field sound based on the HS-PIV data is compared to experimental results. Within both PIV approaches, the typical asymmetric jet flow was detected. However, transient flow field characteristics as high vorticity and maximal velocities peaks could only be observed in HS-PIV data. A strong aeroacoustic source was found immediately downstream of the glottis for both PIV approaches. However, the sources in the jet region could only be observed in the HS-PIV data owing to the averaging procedure for PA-PIV. The comparison between simulated and measured sound spectra revealed good agreement. In contrast, subharmonic tones could not be detected hinting to an additional non-aeroacoustic mechanism of sound generation.

5aSC35. The duration effect in categorical perception of pitch directions and its interactions with vowel quality and language experience. Si Chen (Chinese and Bilingual Studies, The Hong Kong Polytechnic Univ., Hong Kong Polytechnic University, Hong Kong 00000, Hong Kong, qinxi3@gmail.com), Yiqing Zhu, Ratree Wayland, and Edith Kaan (Dept. of Linguist, Univ. of Florida, Gainesville, FL)

It is rarely studied how duration affects categorical perception of pitch directions, and whether this effect interacts with vowel quality or groups of listeners. We created a continuum of fundamental frequency on different vowels and duration values for identification and discrimination tasks. Fifteen native Mandarin speakers and fifteen native English speakers were recruited from the Hong Kong Polytechnic University and the University of Florida. The main effects of vowel quality, groups, pitch directions and duration as well as their interactions significantly contributed to identification rate. The sharpness of the categorical boundary differs significantly when a duration difference reaches 40ms, and it increases first and then decreases as the duration becomes shorter. Tonal listeners exhibit significantly sharper categorical boundary than non-tonal listeners, when a stimuli exceeds 60ms. Moreover, the sensitivity of between- and within-category discrimination differs by groups and duration, whereas vowel quality reaches marginal significance. Duration also affects the peakedness of discrimination significantly. These findings may help explain the observed differences in perceiving segments like vowels and consonants with different length, and shed light on the interaction between the perception of tones and vowel quality due to compensation mechanism in perceiving the intrinsic F0 effect.

5aSC36. Psychometric measurement of vocal strain using a matching task with sustained vowel stimuli. David A. Eddins, Mark D. Skowronski, Supraja Anand (Commun. Sci. & Disord., Univ. of South Florida, 4202 E. Fowler Ave., PCD 1017, Tampa, FL 33620, deddins@usf.edu), and Rahul Shrivastav (Vice President for Instruction, Univ. of Georgia, Athens, GA)

Psychometric measurement of voice quality is important for quantitative assessment of dysphonia. Previous work established the efficacy of a single-variable matching task to index vocal breathiness and roughness using

appropriately designed comparison stimuli, providing a context-independent perceptual task that improves the validity and accuracy of perceptual data compared to magnitude estimation or rating tasks. Physiologically, vocal strain is characterized by excessive vocal fold adduction during phonation, resulting in a glottal pulse with a low duty cycle. The narrowing of the glottal pulse causes the excitation spectrum roll-off to decrease, resulting in a flattening of the spectrum and significant changes in spectral skewness and kurtosis [Moore *et al.*, 1997]. Analogous to comparison stimuli for breathy and rough measurement, a set of noisy sawtooth waveforms was created using modified step filters that varied in spectral slope 1 dB/octave steps, producing a wide range of perceived strain in the comparison stimuli. The comparison stimuli were used in a listening experiment to measure the perceived strain of a set of 15 samples of /a/ selected from the Kay-PENTAX Disordered Voice Database using stratified sampling. The results validate the strain comparison stimuli and extend the matching task paradigm to the three primary dimensions of voice quality perception.

5aSC37. Focus perception in Japanese: Effects of focus location and accent condition. Albert Lee (Dept. of Linguist, Univ. of Hong Kong, Rm. 9.23 Run Run Shaw Tower, Pokfulam HK, Hong Kong, albertlee@hku.hk), Faith Chiu, and Yi Xu (Dept. of Speech, Hearing and Phonetic Sci., Univ. College London, London, United Kingdom)

This study explores the contexts in which native Japanese listeners have difficulty identifying prosodic focus. Theories of intonational phonology, syntax, and phonetics make different predictions as to which focus location would be the most challenging to the native listener. Lexical pitch accent further complicates this picture. In a sentence with mixed pitch accent conditions (e.g., U-A-U), the lexical accent would naturally stand out as more prominent than the unaccented words (U) in terms of modifications to the F0 contour, thus potentially resembling focus. A focus identification task was conducted with 16 native listeners from the Greater Tokyo area. Natural and synthetic stimuli were played to the listeners who then chose which word of the sentence was under focus. Neutral focus (or no focus) was also an option. Stimuli contrasted in accent condition and focus location. Results showed a highly complex interplay between these two factors. For example, accented narrow foci were always more correctly identified (51%) than unaccented ones (28%), whereas the identification rate for final focus was the lowest (31%) among all focus locations. These results are discussed with reference to the research literature on focus production and formal representation of intonation.

5aSC38. Measurement of vocal breathiness perception with a matching task for sustained phonation and running speech. Rahul Shrivastav (Vice President for Instruction, Univ. of Georgia, Athens, GA), Mark D. Skowronski, Supraja Anand, and David A. Eddins (Commun. Sci. & Disord., Univ. of South Florida, 4202 E. Fowler Ave., PCD 1017, Tampa, FL 33620, deddins@usf.edu)

Laboratory measurements of the perception of vocal breathiness typically employ recordings of sustained vowel phonations. Sustained vowels are thought to provide a good representation of the underlying vocal disorder and are relatively easy to produce, analyze, and synthesize. Clinical assessments of voice quality are sometimes based on the production of speech utterances that are more ecologically valid than sustained vowel phonation and that capture the dynamics of laryngeal function, breath support, and a variety of complex cognitive and neuromuscular challenges typically involved in natural speech. As a result, voice quality associated with running speech may correspond more closely with perceived handicap and may represent more relevant treatment targets than sustained vowels. In a listening experiment with naive listeners, breathy voice quality was evaluated for exemplars of sustained /a/ production as well as read speech from a selection of talkers that vary widely in terms of their vocal breathiness. A single-variable matching task was used to index perceived breathiness for both stimulus types. Listener reliability was high, and matching thresholds for the sustained phonation and speech stimuli were highly correlated, demonstrating the efficacy of the matching task for measuring vocal breathiness from speech utterances.

5aSC39. Revisiting English word stress and rhythm in the post-nuclear domain. Jonathan Howell (Montclair State Univ., 1 Normal Ave., Montclair, NJ 07043, howellj@montclair.edu)

In English, grammatical category and word stress are often highly correlated: disyllabic nouns and verbs frequently occur with trochaic and iambic stress, respectively. Kelly and colleagues (1988,1992) suggest that this relationship was shaped by rhythm but may also be subverted by rhythm. We revisit a 1978 production study by Huss, who argued that the usual stress correlation is neutralized in postnuclear position and that acoustic differences are subject to effects of sentence rhythm. Twelve pairs of homophones were embedded in full sentences preceded by a possessive or plural noun with ultimate or penultimate stress: e.g. *farms produce, farm's produce, farmers produce, farmer's produce*. A postnuclear environment was elicited by prepending a contrastive sentence (e.g., *Some say that only supermarket produce may vary in quality but also a farmer's produce may vary in quality*.) Results suggest, contra Huss, that prosodic differences between nouns and verbs persist in the postnuclear environment, despite effects of rhythm. In a mixed-effects logistic regression (N = 18), acoustic cues of syllable duration and maximum intensity were significant predictors of part of speech in nuclear and postnuclear position. We also discuss results of a perception experiment in progress using the production stimuli.

5aSC40. Perception and analysis of "Moe "voice. Sayoko Takano (Kanazawa Institute of Technol., 7-1 Ohgigaoka, Nonoichi 921 - 8501, Japan, tsayoko@neptune.kanazawa-it.ac.jp) and Masashi Yamada (Kanazawa Institue of Technol., K, Ishikawa, Japan)

The term "Moe" is used in Japanese animation and cartoons. "Moe" mainly refers to cute young girls, but it does include direct sexual desire. We investigated the "Moe" voice using "onichan CD" which includes 1200 different utterance from 12 voice actress. "Onichan" means brother in Japanese, and the voice actress says the word with happiness, anger, sadness, adding up to 100 different situations. In experiment one, the impressions of the "Moe" voice was examined using Semantic Differential Method. The subjects answer revealed that Moe is not related to the emotions such as happiness or sadness, but "Moe" is comfort to the listener. Furthermore, "Moe" is mostly associated with the voice actress. This indicates that "Moe" voice is mainly related to the personality. In experiment two, the speech voice was analyzed to measure the acoustic characters. In experiment three, the voice was resynthesized using TANDEM-STRAIGHT. All the results showed that the "Moe" voice is associated with the young girl as small, energetic hypertension.

5aSC41. Cross-cultural differences in the development of emotion perception from face and voice between Japanese and Dutch speakers. Akihiro Tanaka (Psych., Tokyo Woman's Christian Univ., 2-6-1 Zempukuji, Suginami-ku, Tokyo 167-8585, Japan, akihiro@lab.twcu.ac.jp), Disa Sauter (Univ. of Amsterdam, Amsterdam, Netherlands), and Misako Kawahara (Psych., Tokyo Woman's Christian Univ., Tokyo, Japan)

Recent studies have demonstrated cultural differences in multisensory emotion perception from faces and voices. Tanaka *et al.* (2010, Psychol. Sci.) showed that Japanese people are more tuned than Dutch people to vocal processing in adults. The current study investigated how such a cultural difference develops in children and adults. In the experiment, Japanese and Dutch participants observed affective expressions of both Japanese and Dutch actors. A face and a voice, expressing either congruent or incongruent emotions, were presented simultaneously on each trial. Participants judged whether the person is happy or angry. Results in incongruent trials showed that the rate of vocal responses was higher in Japanese than Dutch participants in adults, especially when in-group speakers expressed a happy face with an angry voice. The rate of vocal responses was very low and not significantly different between Japanese and Dutch 5-6-year-olds. However, it increased over age in Japanese participants, while it remained the same in Dutch participants. These results reveal the developmental onset of cultural differences in multisensory emotion perception.

5aSC42. Effects of discrepancy between vocal emotion and the emotional meaning of speech on identifying the speaker's emotions. Sumi Shigeno (Psych., Aoyama Gakuin Univ., 4-4-25 Shibuya, Shibuya-ku, Tokyo 150-8366, Japan, sshigeno@ephs.aoyama.ac.jp)

This study investigates the effects of a discrepancy between vocal emotion and emotion contained in the literal meaning of speech on the identification of speaker's emotions. Except measuring response time, few studies have examined this discrepancy between the emotional meaning of speech and vocal emotion. In the current study, two stimulus conditions were prepared: the "congruent condition," where the emotional meaning of speech was compatible with the tone of voice and the "incongruent condition," where the emotional meaning of speech was discrepant to the vocal emotions. Four native Japanese actors (two males and two females) spoke short sentences emotionally. We used a repeated-measures design with the emotional meaning of speech (happiness/sadness) and the tone of voice (neutral/happy/sad) as two factors. Using a 5-point scale, 31 participants were required to identify the speaker's emotions in a forced-choice task with seven alternatives and report the degree of their confidence for every response. Results indicated that identification of speaker's emotion was successful even in the incongruent condition and that the degree of confidence differed little except the identification of sad voices. The results suggest that Japanese participants infer the speaker's intended emotion even if it is incongruent to the emotional meaning of speech.

5aSC43. Appropriateness of acoustic characteristics on perception of disaster warnings. Naomi Ogasawara (Int. Commun., Gunma Prefectural Women's Univ., 1395-1 Kaminote, Tamamura, Gumma 370-1193, Japan, naomi-o@mail.gpwu.ac.jp), Kenta Ofuji, and Akari Harada (Comput. Sci. & Eng., Univ. of Aizu, Aizuwakamatsu, Japan)

The appropriateness of acoustic characteristics of speech warnings for natural disasters was examined in this study. Three acoustic parameters (voice gender, speech rate, and pitch) were tested in order to find out the best combination of the acoustics which enhances perceived intelligibility, credibility, and urgency by listeners. A hundred native speakers of Japanese listened to short verbal warnings spoken by a male and a female Japanese speaker with normal pitch, 20 Hz higher, or 20 Hz lower pitch; normal speed, 20% faster, or 20% slower speed. Participants rated each stimulus on 1-to-5-scale. A single effect analysis shows that 1) female voice was perceived with higher credibility and urgency; 2) normal pitch was rated best in the three criteria; and 3) normal speed was perceived best except that faster speed was perceived with higher urgency. A mixed effect analysis found that warnings with the combination of female voice, normal speed, and higher pitch most raised credibility and urgency. Faster speed also raised these two criteria although intelligibility was a little sacrificed. Based on the results, we suggest that disaster warnings with female voice and higher pitch would be effective with the appropriate speaking rate to keep good intelligibility.

5aSC44. The relationship between perception of cuteness and duration of voices. Ryohei Ohno (Nihon Univ., 3-25-40, Sacurajosui Setegaya-ku, Tokyo, Japan, Setagaya-ku 156-8550, Japan, ryouhei@kthrlab.jp), Masanori Morise (Univ. of Yamanashi, Kofu, Japan), and Tetsuro Kitahara (Nihon Univ., Tokyo, Japan)

The "cuteness" is one of the most important features in female voices. In fact, voice actresses who are known as cute voice holders are very popular in Japan, particularly for Japanese anime freaks. We have therefore been investigating the relationship between the perception of the cuteness of voices and their acoustic features. Here, we focus on the duration of voices. We hypothesized that it is difficult to perceive cuteness when the voice is short, and we conducted a listening test using voice stimuli with various durations. Given voices (around 1 sec) saying "oniichan" (older brother) by 12 actresses, we extracted 75-ms, 150-ms, 300-ms, and 600-ms parts from them. Participants listened to each stimulus three times and rated their cuteness on a scale of 1 to 5. Experimental results show the following two findings: 1) The stability of ratings for the same voice (i.e., the difference between the highest and lowest ratings for the same voice within each participant) tends to become lower as the voice is shorter. 2) Regarding actresses whose full-version voices are rated 4 or higher by each participant, the participant also rated more than half of the shortened voices (even 75-ms ones) 4 or higher.

5a FRI. AM

5aSC45. Perceptual impressions and acoustical characteristics of simulated "feminine" voices. Maki Waragai and Ayako Shirose (Graduate School of Education, Tokyo Gakugei Univ., 4-1-1 Nukuikita-machi, Koganei, Tokyo 184-8501, Japan, m151423x@st.u-gakugei.ac.jp)

This study aims to clarify characteristics of voice quality that give an impression of femininity. With regard to female voices and gender variation of voices, previous research has studied their acoustical features. Klatt & Klatt (1990) examined male and female voices acoustically and showed that female voices were more breathy than male voices. In order to reveal typical characteristics of female voices, this study investigated perceptual impressions and acoustic features of simulated female voices. Analyzes data were utterances produced by twelve speakers (6 females and 6males, aged 20-22). The speakers were asked to simulate a female "feminine" voice. Firstly, we carried out an impression rating experiment on the obtained data. Twenty-two participants were asked to evaluate the utterances according to various rating scales. The scales included "clear," "thin," "feminine" and so on (Kido and Kasuya, 1999; 2001). Analysis showed that a significant correlation with "feminine" was found in "high," "clear," "thin." Second, /a/ vowels extracted from the utterances were analyzed acoustically. Acoustic parameters were F0, the amplitude gap between H1-H2, and CPP. The relation between these acoustic parameters and the impression ratings of voice quality will be discussed.

5aSC46. Japanese "street seller's voice". Toshiyuki Sadanobu (Graduate School of Intercultural Studies, Kobe Univ., 3-3-18, Nigawa-cho, Nishinomiya-shi 662-0811, Japan, sadanobu@kobe-u.ac.jp), Chunyue Zhu (School of Lang. and Commun., Kobe Univ., Kobe, Japan), Donna Erickson (Kanazawa Medical Univ., Kanazawa, Japan), and Kerrie Obert (Ohio Univ., Athens, OH)

The term "Street Seller's Voice" (henceforth SSV) refers to a category of voice quality which can be uttered only by, but not all, Japanese young girls, especially selling something cute and fashionable. We can hear this voice at cake shops, but never at "wagashi" (i.e., Japanese cake) shops, since "wagashi" is not cute and fashionable, unlike western cake which for Japanese people is cute. In order to catch the nature of SSV we recorded voices of several street sellers, and also conducted MRI experiments with 3 subjects. The results suggest that SSV has a twang quality [e.g., Estill *et al.* 1983, Proc. Stockholm Music Acoustics Conference, 157-174; Honda *et al.* 1995, Vocal Fold Physiology; Voice Quality Control, 23-38], which is manifested in the acoustic signal by, among other things, sustained high energy in the upper frequency regions, and may possibly be produced with high larynx and change in pharyngeal cavity size. One advantage to this type of voice production is it is easily heard in a noisy environment and puts minimum strain on the vocal folds [Estill, personal communication]. [This work was supported by the following grants: JSPS Grants A #15H02605 and A #25240026.]

5aSC47. Effect of temporal fluctuation in speech on perception of humanness of synthesized speech. Fumiya Yokomori, Masanori Morise, and Kenji Ozawa (Graduate School of Medicine and Eng. Sci. Dept. of Education, Univ. of yamanashi, 4-3-11 Takeda, Kofu, Yamanashi 400-8511, Japan, g15mk019@yamanashi.ac.jp)

The sound quality of speech synthesized using modern speech synthesis systems is expected to be approximated to human speech. We investigated the effect of temporal fluctuation of speech on the perception of humanness. Speech stimuli used in the evaluation were generated by voice morphing. The morphing source (morphing rate of 0%) and target (morphing rate of 100%) were speech without temporal fluctuation and original speech, respectively. There were three kinds of morphing factors: fundamental frequency (F0), spectral envelope (SP), and both (F0 + SP). We used MUSHRA as the evaluation method involving two speakers (one male and one female), two phonemes (/a/ and /i/), and two F0s (high and low). Nine morphing rates (every 25% from -100 to 100%) were used, and ten subjects with normal hearing participated in the evaluation. The results show that the stimuli with a morphing rate of 0% scored the lowest humanness for all factors. The stimuli with morphing rates of -100% scored significantly lower than those with a morphing rate of 100% in SP and F0 + SP. The most dominant factor was F0 + SP, and the effect of F0 was the smallest of all factors. [Work supported by JSPS KAKENHI Grant Numbers 15H02726, 16H05899, 16H01734.]

5aSC48. A framework for systematic studies of attitudes in speech. Hiroya Fujisaki (The Univ. of Tokyo, 3-31-12 Ebisu, Tokyo 150-0013, Japan, fujisaki@alum.mit.edu), Sumio Ohno (Tokyo Univ. of Eng., Tokyo, Japan), and Wentao Gu (Nanjing Normal Univ., Nanjing, China)

We present here a framework for a systematic study of attitudes expressed by speech. Although the term "attitude" is commonly used to refer to phenomena at several different levels, we shall assign separate terms to each of these levels for the sake of clarity. The term " attitude" can refer both to a person and an object (either concrete or abstract), here we shall be concerned only with personal relationships. Stances: Those (attitudes) related to official relationships, often influenced by social factors together with long-term personal factors (such as commanding, subordinate, etc.) Attitudes (in the narrow sense): Those referring to private but mid- to long-term relationships, often involving the whole personality (such as kind, friendly, remote, etc.) Manners: Those (attitudes) referring to short-term but somewhat consistent individual behaviors/acts (such as polite, rude, abrupt, etc.) For example, when a teacher tries to call the attention of a young child to an imminent danger, his/her utterance will have commanding stance, kind attitude, and abrupt manner simultaneously. Although in real life these distinctions are not always clearcut, our data collection is based on this principle. The result of a preliminary analysis supports the validity of our approach. [This work was funded by the NSSFC major program 13&ZD189.]

5aSC49. Automatic recognition of negative emotion in speech using support vector machine. Yuto Yamamoto, Masahiro Niitsuma, and Yoichi Yamashita (Graduate School of Information Sci. and Eng., Ritsumeikan Univ., 1-1-1 Noji-higashi, Kusatsu Shiga 525-8577, Kusatsu, Shiga-Pre 525-8577, Japan, yamamoto-ASJ@slp.is.ritsumei.ac.jp)

This paper addresses negative emotion recognition using paralinguistic information in speech for speech dialogue system. Speech conveys not only linguistic information but also paralinguistic and non-linguistic information such as the emotions, attitudes, and intentions. This easily perceivable information plays a key role in a spoken dialog system. However, most of previous speech recognition systems fail to consider this significant information, focusing only on linguistic information, thus hindering the development of more natural speech dialog systems. In order to utilize these significant information for speech dialog systems, this paper focuses on negative emotion recognition from Japanese utterances. 6552-dimensional acoustic features were extracted from 6300 Japanese utterances of 50 people in three emotional state: negative; positive; and neutral. Negative emotion includes anger, sad and dislike. While positive emotion includes favor, joy, and relief. They were classified by SVM and evaluated by a 10-fold cross validation. The experimental result showed the recognition rate of 93.4 % for the classification of negative and positive and 95.0% for the classification of negative and neutral.

5aSC50. Natural language dialog system considering speaker's emotion for open-ended conversation. Takumi Takahashi, Kazuya Mera, Yoshiaki Kurosawa, and Toshiyuki Takezawa (Graduate School of Information Sci., Hiroshima City Univ., 3-4-1, Ozuka-higashi, Asaminami-ku, Hiroshima 731-3194, Japan, takahashi@ls.info.hiroshima-cu.ac.jp)

To respond appropriately to an utterance, human-like communication system, should consider not only words in the utterance but also the speaker's emotion. We thus proposed a natural language dialog system that can estimate the user's emotion from utterances and respond on the basis of the estimated emotion. To estimate a speaker's emotion (positive, negative, or neutral), 384 acoustic features extracted from an utterance are utilized by a Support Vector Machine (SVM). Artificial Intelligence Markup Language (AIML)-based response generating rules are expanded so that the speaker's emotion can be considered as a condition of these rules. Two experiments were carried out to compare impressions of a dialog agent that considered emotion (proposed system) with those of an agent that did not (previous system). In the first experiment, 10 subjects evaluated the impressions after watch four conversation videos (no emotion estimation, correct emotion estimation, inadequate emotion estimation, and imperfect emotion estimation). In the second experiment, another 10 subjects evaluated the impressions after talk with both dialog agents. These experimental results and a demonstration of the proposed system will be shown in the presentation.

5aSC51. Universal vs. language-specific aspects in human vocal attractiveness: An investigation towards Japanese native listeners' perceptual pattern. Anqi Xu and Shue-sum Leung (Dept. of Linguist, The Univ. of Hong Kong, Flat A, 16F, Federate Bldg., No. 292-298 Queen's Rd. West, Sai Ying Pun, Hong Kong 999077, Hong Kong, xuanqi@hku.hk)

Studies on Western societies show that male voices with acoustic parameters encoding a big body size (low F0, narrow formant dispersion, and F0 range) were considered to be attractive, while the opposite was true for female voices (e.g., [Xu et al., 2013, PLoS ONE, 8(4), e62397]). The present work investigates whether Japanese native listeners are guided by the same principles in assessing the voices of the opposite sex. We replicated the design in Xu et al. (2013) with the added parameter of creaky voice, which is prevalent in North America nowadays and hotly debated in terms of attractiveness. Thirty-four heterosexual participants (16 female) rated the attractiveness of synthetic stimuli controlling for F0 height, formant distribution, F0 range and voice quality. Results indicate that their preferences for voice quality are similar with studies on Western societies (breathy > modal > creaky > pressed/tensed). Additionally, low-pitched male voice with narrow formant dispersion and high-pitched female voices were also favorable. Interestingly, a narrow F0 range significantly lowered the attractiveness ratings, regardless of the gender of the voice, which contradicts Xu et al. (2013). These various results are discussed in light of the cross-linguistic/cross-ethnic divergences in vocal attractiveness.

5aSC52. Modeling uncertainty in spontaneous speech: Lexical and acoustic features. Griffin Romigh (Air Force Res. Lab., Dayton, OH), Clayton Rothwell, Brandon Greenwell (Infoscitex Corp., 4027 Colonel Glenn Hwy, Ste. 210, Dayton, OH 45431, crothwell@infoscitex.com), and Meagan Newman (Wright State Univ., Dayton, OH)

Previously, spoken uncertainty has been analyzed using either lexical or acoustic features, but few, if any, studies have used both feature types in combination. Therefore, it is unknown to what extent these feature types provide redundant information. Additionally, prior research has focused on the study of acoustical features of only single words, and it is unclear if those results can generalize to perceived uncertainty in spontaneous speech. The current study elicited spontaneous speech through a team dialogue task in which two people worked together to locate street-level pictures of different houses on an overhead map. The communications were recorded, transcribed, broken into utterances, and presented to 10 individuals who rated each utterance on a 5-pt Likert scale from 1 (very uncertain) to 5 (very certain). A large number of acoustic and lexical features from the literature were calculated for each utterance. Random forest classification (Breiman, 2001) was used to select features and then investigate feature importance individually and also at the aggregate level of feature type. Results indicate that lexical features were much more important than acoustic features and suggest that previous findings using acoustic features might not generalize to spontaneous speech. Additional acoustic features are explored to improve performance.

5aSC53. Assessing the effect of phonetic distance on accommodation. Stephen Tobin (Dept. Linguistik, Universität Potsdam, 406 Babbidge Rd., Storrs, CT 06269-1020, stephen.tobin@uconn.edu) and Adamantios Gafos (Dept. Linguistik, Universität Potsdam, Potsdam, Brandenburg, Germany)

In investigations of phonetic accommodation, convergence is the most frequently reported finding (e.g., Babel, 2012). However, divergence is also attested under some circumstances (Giles, Coupland & Coupland 1991). On the basis of observations and modeling of Tobin and Nam (2009), Tobin, Nam and Fowler (under review) and Kopecz and Schöner (1995), we asked whether and how phonetic distance along some relevant phonetic dimension would modulate accommodation. Following Roon and Gafos (2014) and Gallantucci, Fowler and Goldstein (2009), who report systematic perceptuomotor effects in speech, we used a cue-distractor paradigm to assess the effect of distance. Participants were visually cued to produce a syllable (*ta* or *ka*). 150 ms later they heard a distractor syllable from a 5-step *ta* or *ka* VOT-continuum. Participants were assigned continua either near or distant from their baseline VOT. An ordinal logistic regression of data from 12 participants in the near condition indicated patterns of divergence from the mean distractor VOT among those closest to the distractor, convergence among those farthest, and maintenance at intermediate distances. Pending analysis of data from the distant condition, we conclude that targets must fall within a range neither too close to nor too far from speakers' baseline to induce convergence.

5aSC54. A study on judgement of intoxication using pitch alteration of speech signal in VHF environment. Won-Hee Lee and Myung-Jin Bae (SoongSil Univ., Hyungnam Eng. Buliding 1212, Soongsil University 369 Sangdo-Ro, Dongjak-Gu, Seoul 156-743, South Korea, vbluelovev@ssu.ac.kr)

In this study, speech characteristics from before and after alcohol intoxication has been comparatively analyzed though speech analysis to obtain the degree of intoxication along with its parameters in VHF environment. It is studying speech characteristics of alcohol intoxicated person as pitch and formant and changing of sound levels, but it cannot analyze about before and after alcohol intoxication using these speech characteristics because it is sensitive about environmental changing so we need finding more precise parameters. Especially, speech signal is distorted from the frequency carrier characteristics in VHF environment. Thus, we study speech characteristics from before and after alcohol intoxication using pitch alteration in VHF environment.

5aSC55. Perceived vocal attractiveness by gay listeners in Hong Kong. Shue-sum Leung and Anqi Xu (Univ. of Hong Kong, 2/F, 263 Hang Ha Po, Tai Po, N.T., Hong Kong, Hong Kong, rory@hku.hk)

This study investigates how Cantonese-speaking homosexual males perceive vocal attractiveness; namely, how different phonetic properties of voice influence their judgment. To this end, the present project follows the methodology in Xu et al's (2013) study of vocal attractiveness and emotion in English, and recruited 14 Cantonese-speaking gay men in Hong Kong to rate the attractiveness of digitally synthesised stimuli. The stimuli varied in voice quality (breathy / modal / pressed), fundamental frequency and formant dispersion, all of which were means to project the speaker's intended body size. The attractiveness ratings show that homosexual male subjects prefer male voices with low pitch, breathy voice quality and normal (neither dense nor sparse) distribution of formants. Interestingly, these preferences resemble the ratings by female English subjects in Xu et al's (2013) study with one exception—the females favoured male voices with densely distributed formants, which signal a large body. This divergence needs to be further investigated by follow-up studies targeting different language and sexuality groups, in order to distinguish factors that are universal and language-specific.

5aSC56. Importance of F0 for predicting vocal emotion categorization. Margaret K. Pichora-Fuller, Kate Dupuis (Psych., Univ. of Toronto, 3359 Mississauga Rd., Mississauga, ON L5L 1C6, Canada, k.pichora.fuller@utoronto.ca), and Pascal Van Lieshout (Speech Pathol., Univ. of Toronto, Toronto, ON, Canada)

Affective prosody is used to produce and express specific emotions to conversation partners. While pitch has been identified as a crucial cue for differentiating between emotions, there has been significant variation in the stimuli used by different groups of researchers to examine the acoustic cues necessary for the perception of vocal emotions. The Toronto Emotional Speech Set consists of 2800 items: 200 sentences (carrier phrase "say the word" followed by a target word) spoken by two adult female actors (one younger and one older) to portray seven emotions (anger, disgust, fear, sadness, happiness, pleasant surprise, neutral). In the current study, these stimuli were analyzed to determine which acoustical cues accounted for the most variance in categorizing stimuli into one of the seven pre-determined emotional conditions. The acoustical characteristics of mean duration, mean intensity, mean F0, mean range of intensity, and mean range of F0 were analyzed using a customized Praat script for each of the 2800 stimuli. For both talkers, mean F0 was the most important acoustical cue for accurately categorizing the TESS stimuli into the different emotional conditions. The second most important cue was F0 range for the younger talker and duration for the older talker.

5a FRI. AM

5aSC57. Judgments of American English male talkers who are perceived to sound gay or heterosexual: Certain personality traits are associated with each group of talkers. Erik C. Tracy (Psych., Univ. of North Carolina Pembroke, PO Box 1510, Pembroke, NC 28372, erik.tracy@uncp.edu)

Prior research (Campbell-Kibler, 2011; Gaudio, 1994) discovered that some personality traits (e.g., emotional and intelligent) are associated with gay male talkers and other traits (e.g., masculine and reserved) are associated with heterosexual male talkers. The present investigation examined additional traits listeners associated with four groups of talkers: gay talkers who were perceived as sounding gay, gay talkers who were perceived as sounding heterosexual, heterosexual talkers who were perceived as sounding gay, and heterosexual talkers who were perceived as sounding heterosexual (Tracy, Bainter, & Satariano, 2015). After hearing a spoken utterance, listeners rated talkers along eight traits: bored, confident, intelligent, mad, old, outgoing, sad, and stuck-up (Tracy & Charlton, 2016). During the experiment, the talkers' self-identified and perceived sexual orientation were not referenced. The results demonstrated that the gay and heterosexual talkers who were perceived as sounding gay were rated as being significantly more confident, mad, stuck-up, and outgoing; the gay and heterosexual talkers who were perceived as sounding heterosexual were rated as being significantly more bored, sad, and older. There were no significant rating differences for intelligence. It was concluded that a talker's perceived sexual orientation, rather than their self-identified sexual orientation, influenced which personality traits listeners associated with them.

5aSC58. Phonological structure of toasting practices in Japan. Yoko Sakamoto (Premedical Sci., Dokkyo Medical Univ., 880 Kitakobayashi, Mibu, Shimotsugagun, Tochigi 3210293, Japan, y-saka@dokkyomed.ac.jp)

The aim of this study was to investigate toasting practices in Japan from the point of view of phonological structure. In Japan, toasting practices means calls or songs to encourage someone to drink alcohol at one stroke. This habit has lasted for 30 years among university students and company young employees in Japan, although many campaigns have been held against toasting practices. The background of this habit may have complex factors, such as social, psychological, etc., but one of the factors may be in the calls and songs themselves. Thus, in the present study, calls of songs of toasting practices were collected as database, and the phonological structure (foot, syllable, mora, phoneme) and the tempo was analyzed. The result showed that the calls and songs of toasting practices in Japan are categorized into three groups; calls, parody songs and others (such as famous phrases), but these calls and songs have common phonological structure. They tend to have 4 morae 2 feet structures and combinations of Motherese syllable structures. The tempos of calls and songs were 120-140, which were the double speed of adult's normal heartbeat. Thus, it has been concluded that toasting practices in Japan have regularity in phonological structure.

5aSC59. Constructing boyishness: A sociophonetic analysis of Japanese anime voice actresses performing male roles. Ryan C. Redmond (Linguist, Univ. of California Davis, 2505 5th St. Apt. 117, Davis, CA 95618, rcredmond@ucdavis.edu)

The present study investigates the performance of different gendered roles by Japanese female seiyuu [voice actors] through analysis of fundamental frequency (f0). Ten voice actresses who have been known for commonly voicing male characters were sampled for analysis. For each actress, two characters from their repertoire were chosen (one male, one female), and from each character 75 intonation phrases were sampled, for a total of 1500 utterances. For each of these samples, pitch data were extracted and analyzed. Results for both male and female role performances display an average f0 that generally falls inside the category previously defined as hegemonically "feminine" in Japanese. While male role measurements were on average lower than female role measurements, only occasionally did any actress produce pitches in the average expected range for a Japanese man. However, because the female role measurements were significantly higher than that of average Japanese women, the differences between the "male role" and "female role" categories proved statistically significant in terms of minimum pitch, maximum pitch, mean pitch, and pitch range. A theory is proposed that the hyper-femininity of many female roles is what necessitates this difference.

Session 5aSP

Signal Processing in Acoustics: General Topics in Signal Processing II

Ning Xiang, Chair
School of Architecture, Rensselaer Polytechnic Institute, Greene Building, 110 8th Street, Troy, NY 12180

Contributed Papers

7:55

5aSP1. Sound quality improvement based on weighted AM/FM modulations for parametric loudspeaker. Kirara Ariyoshi, Takuya Kimura, Takahiro Fukumori, Masato Nakayama, and Takanobu Nishiura (Ritsumeikan Univ., 1-1-1, Noji-higashi, Kusatsu, Shiga 525-8577, Japan, is0152hv@ed.ritsumei.ac.jp)

A parametric loudspeaker has high-directivity, and it utilizes a modulated wave that is designed by modulating an ultrasound wave with an audible sound wave. The conventional modulation methods, amplitude modulation (AM) and frequency modulation (FM), have been proposed. Furthermore, an AM/FM method has been proposed that can improve the sound pressure level (SPL). However, its sound quality is inferior to the AM method. The AM/FM method utilizes constant modulation parameters in all frequency bands. However, the power of audible sound demodulated by AM and FM modulations is different at each frequency band. Therefore, we propose a new AM/FM method that utilizes the weighted modulation parameters in each frequency band and can improve the sound quality. The proposed method has comparable sound quality to the AM method and a comparable SPL to the FM method. In this proposed method, we divided all frequency bands into three frequency bands and utilized the optimum weighted modulation parameters in the three frequency bands on the basis of the results of the preliminary experiments. As a result of evaluation experiments, we confirmed the effectiveness of the proposed method compared with the conventional methods.

8:10

5aSP2. MATLAB visualization program for analysis of C3D motion capture data of speech and language. Andrew Barnes, Rebecca Mental, Brooke Macnamara, and Jennell Vick (Dept. of Psychol. Sci., Case Western Reserve Univ., 11635 Euclid Ave., Rm. 308, Cleveland, OH 44106, acb48@case.edu)

Investigation of the underlying physiology of speech and language motor control often includes analyses of kinematic parameters derived from fleshpoint tracking. A MATLAB program has been developed to visualize, boundary mark, and analyze aspects of motion capture data in a C3D file format. C3D is the biomechanics data standard for binary 3D data that is widely used in many industries and motion research. Included in the format is marker names, arrays of position data, and analog inputs. While this format is highly versatile, it is difficult to assess the comparison of the markers positions without a visualization program. The developed program allows for flexible comparison of any fleshpoint markers, in any set of repetitions. It also allows for easy visualization of the differences between varying data conditions. Data analysis is done on-the-fly for information such as joint angles through three markers, distance between any two markers, and variability analyses of any selected data. Data are output in a format suitable for additional statistical analyses. Recent analyses using this program were completed on highly complex data streams with over 30 fleshpoint markers. The program provides an incredibly useful tool for exploration of these data and quick analyses of coordination and variability.

8:25

5aSP3. Development of a 3D impulse response interpretation algorithm. Federico N. Cacavelos, Alejandro Bidondo, and Augusto Bonelli Toro (Engineer, Universidad Nacional de Tres de Febrero, Mosconi 2736, CABA, Saenz Peña 1674, Argentina, fnahuelc@gmail.com)

The development of software 3D interpretation impulse responses is presented. This tool allows to identify the spatial provenance for time windows of an impulse response recorded in a soundfield microphone. In this way, it is possible to discern between the direct sound and reflections both temporarily and spatially. To understand the basics of this method, a brief introduction to all concepts involved is presented. Alternatively, different types of procedures are shown for data evaluation. Finally, system applications are shown in controlled acoustical environments so evaluating the precision and the potential of this tool.

8:40

5aSP4. The acoustical features analysis of Indo-Pacific humpback dolphins' whistle and modeling research. Siyuan Cang, Xueli Sheng, Jintao Sun, Dian Lu, Jingwei Yin, and Longxiang Guo (Harbin Eng. Univ., College of Underwater Acoust. Eng., Harbin Eng. University, Harbin, Heilongjiang 150001, China, cangsiyuan@hrbeu.edu.cn)

Depending on species-specific characteristics of echolocation and recognition, dolphin could travel in groups without collision, and they could also prey and defend mutually. As a multi-harmonics, frequency-modulated signal, the variety of whistles revealed the rich group identification and location information of free-ranging Indo-Pacific humpback dolphins (Sousa chinensis). This paper introduces the features of whistle on the basis of Joint time-frequency distribution, dealing with recording data collected in the area of Guangxi and Zhanjiang seacoast, in China. Then the ability of individual identification is discussed from the correlation of whistles, we will get the knowledge of dolphins how to cognize the communication information in the complex environment. The transmission loss function will be depicted using the BELLHOP at last. By extracting the acoustical features, the modeling of whistle signal is implemented based on the Sinusoids Model (Buck, 2000). Specifically, the validity of mathematical model is proved according to comparison with the recording data.

8:55

5aSP5. Comparisons of conventional and frequency-difference beamforming. Alexander S. Douglass and David R. Dowling (Mech. Eng., Univ. of Michigan, 2010 AL, 1231 Beal, Ann Arbor, MI 48109, asdougl@umich.edu)

Conventional beamforming can isolate ray arrival directions from transducer array recordings, provided the array element spacing does not significantly exceed half of a signal wavelength. However, when the signal frequency is high and the array elements are many wavelengths apart, conventional beamforming may break down due to spatial aliasing. Frequency difference beamforming avoids this limitation by utilizing a quadratic product of complex signal amplitudes at different frequencies to extract wave propagation information at a difference frequency chosen low enough to avoid aliasing. For this presentation, the performance of frequency

3403   J. Acoust. Soc. Am., Vol. 140, No. 4, Pt. 2, October 2016

5th Joint Meeting ASA/ASJ   **3403**

difference beamforming was evaluated two ways. First, simulations of acoustic propagation in a 106-m-deep shallow-ocean sound channel using the BELLHOP ray-tracing code were completed with frequency-sweep signals from 11 kHz to 33 kHz and 1100 Hz to 3300 Hz broadcast from a single source to a 56-m-long 16-element vertical array. Second, lab experiments in a 1.07-m-diameter 0.8-m-deep cylindrical water tank were conducted with pulse signals having center-frequencies of 165 kHz and 15 kHz broadcast from a single projector to a 0.7-m-long 14-element horizontal array. In both cases, frequency difference beamforming utilizing the higher frequency signal produced results comparable to conventional methods applied to the lower frequency signal. [Sponsored by ONR.]

9:10

5aSP6. A novel method to detect rising and setting tones. Suci Dwijayanti and Masato Miyoshi (Graduate School of Natural Sci. and Technol., Kanazawa Univ., Kanazawa, Ishikawa 920-1192, Japan, dwijas@stu.kanazawa-u.ac.jp)

Rising or setting tone is a natural phenomenon caused by the chorus emission, which has a discrete frequency time structure and increasing or decreasing frequencies. It usually happens in the Very Low Frequency (VLF) range and sounds like bird tweeting. Since the chorus emission seems to have a clue to space research for safe and secure satellite operations, there is a strong demand to detect such tweeting tone. To this end, we develop a new signal processing tool wherein the running spectra of a tone are first treated as the signals received with individual sensors in a linear senorarray. Then, the estimation of signal parameter via rotational invariance techniques (ESPRIT) algorithm is applied to estimate the rising or setting speed of the tone. In order to evaluate the proposed tool, we conduct simulations using chirp signals of different rising or setting speeds. The simulation results show that the rising or setting speed of a tone can accurately be estimated with the proposed tool.

9:25

5aSP7. An analysis of methods for the acoustic localization of unattended aerial vehicles. David Grasing (ARDC U.S. Army, Bldg. 407, Picatinny, NJ 07806, david.b.grasing.civ@mail.mil)

In this paper methods, models, and difficulties associated with the acoustical detection and tracking of Unattended Aerial Vehicles (UAVs) are presented. The chief advantages of acoustics include passive sensing, the ability to operate in day or night, and relatively low cost. Acoustics also presents several notable challenges including low signal to noise, and issues arising from ground reflections. It is the presence of ground reflections that is explored in particular detail as it presents one of the more unique challenges to this problem. Issues arise because the differences in path length between the line-of-sight and ground reflection arrival of the source are of the same order as the wavelengths of the lower frequencies we are able to detect. Methods for bearing estimation will also be presented, explored, and evaluated in terms of their efficacy. The harmonic structure created by the UAV's propellers and engines are extracted. Bearings are estimated using various beamforming methods including CB, MVDR, and MUSIC. A black box optimization method based on the Nelder-Mead method on the surface of sphere is presented and used to evaluate different bearing estimation criteria. Finally, bearing estimates at different frequencies are combined via coherent and incoherent methods.

9:40–10:00 Break

10:00

5aSP8. Application of scanning acoustic and photoacoustic microscopes in orthopedic surgery. Yoshihiro Hagiwara (Dept. of Orthopaedic Surgery, Tohoku Univ. School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai 980-8574, Japan, hagi@med.tohoku.ac.jp) and Yoshifumi Sajo (Dept. of Biomedical Imaging, Graduate School of Biomedical Eng., Tohoku Univ., Sendai, Japan)

Demands for orthopedic surgery, decreasing pain and disabilities, is increasing in developed countries with older nations. Together with increasing number of patients, its cost is now rising. Among modalities for noninvasive diagnosing techniques, such as plain X-rays, computed tomography,

magnetic resonance imaging (MRI) and ultrasonography (US), MRI is a powerful tool for evaluating anatomical abnormalities. However, location and economic limitations for routine use remain. US is quick and inexpensive with a higher resolution than MRI. US plays an important role in assessing musculoskeletal soft tissues. A scanning acoustic microscope (SAM) characterizes biological tissues by estimating the elastic parameters based on sound speed. Biomedical photoacoustic (PA) imaging has the unique capability of combining high optical contrast and high ultrasound resolution in a single modality. Osteoarthritis and frozen shoulder are the major problems in musculoskeletal diseases. Osteoarthritis is a degenerative joint disorder characterized by the progressive degeneration of articular cartilage, osteophyte formation, and subsequent joint space narrowing. Further, frozen shoulder is characterized with severe pain and a decrease in shoulder motion, which is caused by joint capsular stiffness. We applied SAM and PA to animal models for assessing osteoarthritis and human samples with frozen shoulder.

10:15

5aSP9. Convolutional bidirectional long short-term memory hidden Markov model hybrid system for polyphonic sound event detection. Tomoki Hayashi (Nagoya Univ., Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan, hayashi.tomoki@g.sp.m.is.nagoya-u.ac.jp), Shinji Watanabe (Mitsubishi Electric Res. Labs. (MERL), Cambridge, MA), Tomoki Toda (Nagoya Univ., Nagoya, Japan), Takaaki Tori, Jonathan L. Roux (Mitsubishi Electric Res. Labs. (MERL), Cambridge, MA), and Kazuya Takeda (Nagoya Univ., Nagoya, Japan)

In this study, we propose a polyphonic sound event detection method based on a hybrid system of Convolutional Bidirectional Long Short-Term Memory Recurrent Neural Network and Hidden Markov Model (CBLSTM-HMM). Inspired by the state-of-the-art approach to integrating neural networks to HMM in speech recognition, the proposed method develops the hybrid system using CBLSTM to estimate the HMM state output probability, making it possible to model sequential data while handling its duration change. The proposed hybrid system is capable of detecting a segment of each sound event without post-processing, such as a smoothing process of detection results over multiple frames, usually required in the frame-wise detection methods. Moreover, we can easily apply it to a multi-label classification problem to achieve polyphonic sound event detection. We conduct experimental evaluations using the DCASE2016 task two dataset to compare the performance of the proposed method to that of the conventional methods, such as non-negative matrix factorization (NMF) and standard BLSTM-RNN, and also investigate the effect of network structures on the detection performance.

10:30

5aSP10. Chatter diagnosis using Mel frequency cepstral coefficient of vibrational signal for various operating conditions. KwangHun Jeong, JeWon Lee, and Junhong Park (Mech. Eng., Hanyang Univ., 222, Wangsimni-ro, Seongdong-gu, Eng. Ctr., 306, Seoul 04763, South Korea, zjavbfk@hanyang.ac.kr)

The chatter reduces quality of the machined parts and induces vibration and sound that is different from those under normal cutting process. A variety of cutting conditions are significantly correlated with the occurrence of chatter. After changing of the cutting conditions such as rotational speeds and depths of machine tools blade, the vibration signals was measured using vibrational accelerometers. The chatter diagnosis algorithm using Mel Frequency Cepstral Coefficient (MFCC) and Deep belief network (DBN), one of the Deep learning algorithm is proposed. Deep belief network is effective on classifying complicated signals as it represents the hierarchical cognitive process of a human brain. To acquire features from the distinctive noise and vibration under normal and chatter status, the MFCCs was used. Chatter diagnosis algorithm, DBN using MFCCs as input features, is suggested and verified.

10:45

5aSP11. Application of acoustic contrast control for out door music performance: Feasibility study and experiments. Yang-Hann Kim (KAIST, Dept. of M.E., Sci. Town, Daejon-shi 305-703, South Korea, yanghann@kaist.ac.kr), Wanjung Kim, Hwan Kim, Jueseok Kim, Sang-Hyeon Kim (SQand, Daejeon, South Korea), Jungwoo Choi (KAIST, Daejeon, South Korea), and Jong-Hwa Lee (SQand, Daejeon, South Korea)

Ideal outdoor music hall will be what has invisible but acoustically hard wall. This rather impossible objective can be achieved, at least, in some extent, for example allowing quite acceptable sound level in certain zone or zones where most residence lives and hear unwanted sound. This challenging problem can be tackled by finding the solutions that determines the magnitudes and phases of speakers, which surround the ideal hall's wall, that minimizes acoustic contrast between the listeners zone (bright zone) and residence zone (dark zone). It is interesting to find that acoustic contrast control essentially builds up acoustically hard wall, in other words, it makes large impedance mismatch on the wall. This bright observation, however, has been essentially based on two dimensional simulation and experiment. Next step has to go how well this can be extended to general 3D cases. We attempted to have another layer of speaker arrays, and requiring bright zone inside of the double array with certain height, therefore making a cylindrical volume of bright zone, and dark zone having certain angle, between two lines: one connects between the center and the bottom array and another one is the line between the center and upper array, rotating these two lines makes a diverging volume outside of the wall that has certain angle. Simulations and experiments demonstrated how well it can be practically applied and what are its limitations.

11:00

5aSP12. Modified DCTNet for audio signals classification. Yin Xian, Yunchen Pu, Zhe Gan (Duke Univ., 111 Blue Crest Ln., Durham, NC 27705, poline3939@gmail.com), Liang Lu (Univ. of Edinburgh, Edinburgh, United Kingdom), and Andrew Thompson (Oxford Univ., Oxford, United Kingdom)

We introduced the use of modified DCTNet for audio signals feature extraction. The modified DCTNet is a development of DCTNet, with its center frequencies of filterbanks geometrically spaced. The modified DCTNet is adaptive to different acoustic scales, and it can better capture low frequency acoustic information that is sensitive to human audio perception. We use features extracted by the modified DCTNet and put them to classifiers. Experimental results on bird song classification, music genre classification, and artist identification show that the modified DCTNet and RNN improve classification rate, and achieve state of the art performance.

11:15

5aSP13. A nested beamforming approach to sound source separation using an array-of-arrays recording system. Jorge A. Trevino Lopez, Shuichi Sakamoto, and Yôiti Suzuki (Graduate School of Information Sci. and Res. Inst. of Elec. Commun., Tohoku Univ., 2-1-1 Katahira, Aoba-ku, Sendai, Miyagi 9808577, Japan, jorge@ais.riec.tohoku.ac.jp)

Beamforming is a technique used with microphone arrays to selectively isolate sounds according to their direction of incidence. However, conventional methods struggle to separate sound sources if the spatial resolution of the recording system is not enough to define a recording zone containing only the target. In this presentation, a sound source separation method using multiple beamforming arrays is introduced. The proposal consists of three stages. The first stage applies modal beamforming methods to individual spherical microphone arrays; this results in a virtual array of directional microphones. The second stage distributes the recordings from the virtual microphones onto a new angular coordinate system. This stage transforms the radial coordinates for the different sound sources into pairs of parallax angles between them and the spherical microphone arrays. Finally, the third stage uses a least-squares optimized beamforming method defined on the parallax-angle coordinate system. The combination of these three stages takes advantage of parallax differences between spherical microphone arrays and allows for the isolation of sound sources at specified positions. The new method is evaluated using a computer simulation of two 64-channel spherical microphone arrays. The proposal shows satisfactory results even in the presence of aligned interfering sources, which hinder separation using conventional beamforming methods.

11:30

5aSP14. Sound enhancement system using selective binary filtering. Tomomi Suzuki (Graduate School of Information Sci., Nagoya Univ., Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan, suzuki.tomomi@g.sp.m.is.nagoya-u.ac.jp), Takanori Nishino (Graduate School of Eng., Mie Univ., Tsu, Mie, Japan), Yoshio Ishigiro (Inst. of Innovation for Future Society, Nagoya Univ., Nagoya, Aichi, Japan), and Kazuya Takeda (Graduate School of Information Sci., Nagoya Univ., Nagoya, Aichi, Japan)

Our proposed system allows users to selectively focus on particular sounds by choosing from various filtering options. Our goal is to develop a sound enhancement system which increases human auditory performance, especially through selective listening. To manage multiple filtering options numerically, we propose a series of listening patterns consisting of sets of binary indicators which show whether or not signals from a particular source will be presented to the listener. We then calculate the similarity of each pattern and order these patterns on the basis of their similarity. We conducted an experiment to evaluate our proposed method. Subjects listened to a sound recording using the various patterns in the order determined using the proposed method, as well as listening in an order determined using a comparative method. Subjects were asked to evaluate the naturalness of the transitions from pattern to pattern using each method, to select a listening pattern corresponding to one which was described to them, and to evaluate ease of use when searching for a desired listening pattern. Results support the feasibility of the proposed sound enhancement system. One issue which needs to be resolved is how to deal with complex environments with numerous sounds as possible targets.

Session 5aUWa

Underwater Acoustics: Underwater Acoustic Propagation

Ahmad T. Abawi, Chair
HLS Research, 3366 North Torrey Pines Court, Suite 310, La Jolla, CA 92037

Contributed Papers

8:00

5aUWa1. Long range propagation and refraction of acoustic signals measured on the Comprehensive Nuclear Test Ban Treaty hydroacoustic network in the Indian Ocean. David R. Dall'Osto (Acoust., Appl. Phys. Lab. at Univ. of Washington, 1013 N 40th St., Seattle, WA 98105, dallosto@apl.washington.edu)

This paper investigates 3 days data collected on the Comprehensive Nuclear Test Ban Treaty hydro-acoustic network in the Indian Ocean. This data includes that from March 8, 2014, the day the Malaysian airliner MH370 went missing—presumably crashing into the Indian Ocean as evidenced by found aircraft debris. This presentation focuses on the earthquakes and other anthropogenic and natural sounds recorded at South Station of Diego Garcia [7.662 S, 72.613 E] and Cape Leeuwin located ~5,000 km away off of SW Australia [34.981 S, 114.144 E]. These measurements show that "hydroacoustic blockage," shadowing caused by a land mass located between the source and the receiving station, is incomplete and that acoustic signals can be detected on the far side of an island. The land mass significantly alters the time and frequency dependence of the signal, in comparison to an unblocked path. Acoustic models are used to investigate this propagation phenomena, and suggest capability for localizing sound sources using "blocked" stations. The presentation also includes analysis of a conspicuous signal, localized to an area in the West Indian Ocean. The source of this signal is unknown, although it originated shortly after the last satellite last transmission from the missing airliner.

8:15

5aUWa2. Bottom-diffracted surface-reflected arrivals in the North Pacific. Ralph A. Stephen (Woods Hole Oceanographic Inst., 360 Woods Hole Rd., Woods Hole, MA 02543-1592, rstephen@whoi.edu), Peter F. Worcester (Scripps Inst. of Oceanogr., La Jolla, CA), S. T. Bolmer (Woods Hole Oceanographic Inst., Woods Hole, MA), Ilya A. Udovydchenkov (The MITRE Corp., Bedford, MA), and Matthew A. Dzieciuch (Scripps Inst. of Oceanogr., La Jolla, CA)

Bottom-diffracted surface-reflected (BDSR) arrivals were first identified in the 2004 Long-range Ocean Acoustic Propagation Experiment (LOA-PEX) in the North Pacific (Stephen et al., 2013, JASA, **134**, 3307-3317). In deep water, ambient noise and PE predicted arrivals are sufficiently quiet that BDSR arrivals, scattered from small seamounts, can be the largest amplitude arrivals observed. The Ocean Bottom Seismometer Augmentation in the North Pacific (OBSANP) Experiment in June-July 2013 studied BDSRs at the same site in detail. They are most readily identified by their move-out on lines of transmissions and are clearest on the vertical component channel. There appear to be two classes of diffractors. The diffraction point for Type 1 BDSRs occurs on the side of small seamounts and is substantially out of the source-receiver sagittal plane (for example, the BDSR observed on NPAL04). The diffraction point for Type 2 BDSRs is essentially in the source-receiver sagittal plane and occurs on the relatively featureless abyssal plain. In at least one case the same BDSR is observed for 77.5, 155 and 310 Hz M-sequence transmissions. Other characteristics of BDSRs will be discussed. [Work supported by ONR.]

8:30

5aUWa3. ACOustic keyhole. Rhett Butler, Frederick K. Duennebier, and Bruce M. Howe (School of Ocean and Earth Sci. and Technol., Univ. of Hawaii at Manoa, 1680 East-West Rd., Post 602, Honolulu, HI 96822, rgb@hawaii.edu)

The Aloha Cabled Observatory (ACO) is located on the seafloor about 100 km north of O'ahu at 4800 m depth. Seismoacoustic T-waves generated by earthquakes in the South Pacific from Tonga-Kermadec to the Solomon Islands travel 40-70°Δ, passing through an acoustic keyhole—the Ka'ie'ie Waho channel—between O'ahu and Kaua'i before reaching ACO. In contrast to ACO and the Pacific seafloor at depths greater than 4500 m, the channel shallows to 1 and 3 km in depth, blocking deeper SOFAR ray/mode paths. T-waves from >50 earthquakes (Mw 6.1 to 8) are reviewed, correcting apparent velocities for near-source P-wave path effects. We examine T-waves crossing the Ka'ie'ie Waho channel and descending to ACO, comparing and contrasting with T-waves propagating unhindered across the North Pacific from events in offshore Japan (Tohoku aftershocks), Kuril Islands, Aleutian Islands, and offshore Alaska. Preliminary review indicates that the T-waves exhibit seismoacoustic modal-coupling to the seafloor, exhibiting a structure observed previously* at the Hawai'i-2 Observatory (H2O) at 5 km depth between Hawai'i and California. *R. Butler (2006), J. Acoust. Soc. Am. **120**(6), 3599-3606. *R. Butler and C. Lomnitz (2002), Geophys. Res. Lett. **29**(10), 57 1-4.

8:45

5aUWa4. Acoustic features of sharp and continuous three-dimensional curved coastal fronts. Brendan J. DeCourcy (Mathematical Sci., Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180, decoub@rpi.edu), Ying-Tsong Lin (Appl. Ocean Phys. & Eng., Woods Hole Oceanographic Inst., Woods Hole, MA), and William L. Siegmann (Mathematical Sci., Rensselaer Polytechnic Inst., Troy, NY)

Three-dimensional curved shelf-slope fronts have been modeled with sharp [Lin and Lynch, J. Acoust. Soc. Am. (2012)] and continuous [DeCourcy et al., J. Acoust. Soc. Am. (2016)] sound speed changes. For the sharp front, expressions were found for determining dependence of acoustic quantities such as modal speeds on feature-model parameters. Asymptotic approximations are used to simplify terms in the dispersion relation for the horizontal wavenumber, and further analysis leads to convenient and accurate formulas for the parameter dependence. Corresponding results will be presented for the continuous front, which possesses a different class of acoustic modes known as tunneling modes that carry larger amounts of energy across and along the front than other modes. Results for both frontal models will be discussed, comparing and contrasting characteristics including wavenumber distributions, mode types, and parameter dependences. An important goal is to examine the similarities and differences for sharp or continuous fronts in the sensitivity of acoustic quantities to feature parameters, such as front location and width, bottom slope angle, and source frequency. This would permit identification of situations where the more convenient sharp front model is appropriate. [Work supported by ONR Grants N00014-14-1-0372 and N00014-11-1-0701.]

5aUWa5. Analysis of sound propagation in the front region off the east of Tsugaru straits. Yoshiaki Tsurugaya (Sanyo P. T., 11-14 Ooi 5-Chome, Shinagawa-Pref., Tokyo 140-0014, Japan, tsul@mvb.biglobe.ne.jp), Toshiaki Kikuchi (National Defence Acad., Ooyabe, Kanagawa-pref., Japan), and Koichi Mizutani (Univ. of Tsukuba, Tsukuba, Ibaraki-pref., Japan)

Off the east of Tsugaru straits, located in the north of the main-island, there are three big water masses. One is the Oyashio (Kuril current) that is a cold water mass. Other water masses are the warm water masses: the Tsugaru warm current, and Kurosio extension. And, the Tsugaru warm eddy is made from the Tsugaru warm current on the Oyashio. Therefore, the front is generated in the region to which those water masses bound. The temperature structure changes rapidly in that region. In addition, the bottom depth is changed too. Then, bi-directional sound propagation is analyzed on the line of a right-angled direction to the front that Tsugaru warm eddy is bounded to the Oyashio. It is considered about the sound propagation of 220 km that passes the front from the Tsugaru warm eddy to the Oyashio cold water mass, and the opposite path. In the front region, the sound propagation is changed, and suffers the impact of the bottom at the same time. Therefore, the propagation configuration is different in the bi-directional propagation. The transmission loss in the Tsugaru warm eddy was large when the sound source crosses the front from the Oyashio.

5aUWa6. Influence of source frequency spreading on transmission loss patterns. Elisabeth M. Brown (Mathematical Sci., Rensselaer Polytechnic Inst., Troy, NY 12180, browne6@rpi.edu), Allan D. Pierce (Boston Univ., East Sandwich, MA), and William L. Siegmann (Mathematical Sci., Rensselaer Polytechnic Inst., Troy, NY)

A classical experiment [Wood and Weston, Acustica 14 (1964)], conducted in Emsworth Harbor at low tide, examined propagation in a 1 m mud layer overlying gravel. Hydrophones on the mud bottom transmitted six source frequencies between 4 and 72 kHz and received signals at ranges up to 50 m. Signal levels show modal fluctuations at lower frequencies but not at higher frequencies, even though the authors did not perform data smoothing. This behavior is hypothesized to arise from a frequency-dependent spread of transmitted frequencies. A Pekeris two-layer model of mud and gravel with a pressure release mud surface is used to idealize the waveguide. Transmission loss is calculated using a weighted frequency average of time-averaged intensity, with a Lorentzian function (Cauchy distribution) modeling the frequency spread. Appropriate values for the Q-factor of the distribution are examined. Compressional wave attenuation values from a recent inversion of the data [Pierce, *et al.*, POMA, accepted (2015)] are incorporated into the calculations. Computational results demonstrate the same behavior in intensity fluctuations as in the data with identifiable modal interference at the three lower frequencies and smoothed patterns at the three higher. [Work supported by the ONR.]

5aUWa7. Analytical formulas for incoherent transmission loss in shallow water based on effective approximations of seafloor depth and reflectivity. Zhi Y. Zhang (Defence Sci. and Technol. Group, West Ave., DSTG/MD, Edinburgh, SA 5111, Australia, YongZhang.work@gmail.com)

We derive analytical formulas for modelling incoherently averaged transmission losses in shallow water environments by taking into account the effects of energy spreading, in-water absorption, sea surface scattering, and leakage into the seafloor. The formulas are based on the effective depth approximation to include the effect of sound penetration into the seafloor and use effective two-parameter representations of seafloor reflectivity. Comparison with published measurement data at several shallow water sites worldwide shows that the formulas are capable of reproducing the main features of shallow water sound propagation in environments with different types of sediments. The formulas are useful for quick assessments of incoherent transmission losses in generic yet representative underwater environments where detailed environmental knowledge is lacking or the fine details of underwater sound propagation is not required.

5aUWa8. Modeling surface scatter in the channel impulse response with a Helmholtz-Kirchhoff integral. Edward Richards, William Hodgkiss, and Heechun Song (Scripps Inst. of Oceanogr., Univ. of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093, edwardlrichards@ucsd.edu)

Several experimental measurements of channel impulse responses (CIRs) using underwater communication bandwidths have shown distinct striation features arriving up to 10 milliseconds after the main surface bounce, e.g. Badiely *et al.* [J. Acoust. Soc. Am. **132**, EL290 (2012)]. Choo *et al.* [J. Acoust. Soc. Am. **136**, 1046 (2014)] used an intuitive geometric ray model to argue that the source of these arrivals was the reflection from a wave trough near the receiver. This presentation extends Choo's approach and focuses on a quadrature solution of the time domain Helmholtz-Kirchhoff (HK) integral in a half space. This approach provides the same intuitive geometric ray perspective on the surface scattering, but additionally models arrival amplitudes near caustics and higher order catastrophe phenomena induced by surface waves. It is shown that a simple shadowing correction of the HK result produces all of the major features of the full Helmholtz scattering integral solution. The HK integral is computationally tractable, accurate, and can show the path from the surface that contributes to each CIR arrival.

5aUWa9. Random matrix theory simulation of propagation through internal waves. Kathleen E. Wage (Elec. and Comput. Eng., George Mason Univ., 4400 University Dr., MSN 1G5, Fairfax, VA 22030, kwage@gmu.edu) and Lora J. Van Uffelen (Ocean Eng., Univ. of Rhode Island, Narragansett, RI)

Hegewisch and Tomsovic proposed using random matrix theory (RMT) to simulate the effect of internal wave scattering on long-range propagation in deep water [JASA, 2013]. They model mode propagation as the multiplication of a series of unitary propagator matrices whose statistics are derived using perturbation theory. A simplifying assumption facilitates the generation of broadband time series. One potential advantage of the RMT approach is that it requires fewer computations than a full wave simulation to produce a large set of statistical realizations of the received time series [Hegewisch & Tomosovic, Europhys. Lett. 2012]. Efficient simulation techniques are especially important for applications that require predictions of the timefront and/or its statistics at a large number of possible receiver ranges. For example, simulating data sets for acoustic gliders operating within a tomographic array is computationally intensive. While Hegewisch and Tomsovic illustrated the performance of their approach for one deep water simulation environment, they also highlighted the need for additional research. This talk investigates the feasibility of using RMT to efficiently simulate broadband time series for use in analysis of glider data. The underlying assumptions of the RMT approach will be assessed based on results of prior experimental work. [Work sponsored by ONR.]

5aUWa10. The optimal weight vectors research on the far-field sound field control of the low-frequency phased array. Haifeng Lee (Inst. of Acoust., CAS, No. 21,West Rd. of Beishihuang, Haidian, Beijing 100190, China, cust77@mail.ioa.ac.cn) and Li Ma (Inst. of Acoust., CAS, Peking, China)

The theory and simulation studies were developed to analyze the far-field sound field control effect of the low-frequency phased array emitting the different weight vectors. First, the relationship between the weight vector and the mean square deviation of time-delay estimation were set up theoretically. Second, using the relationship, the optimal weight vector was solved and discussed. Finally, under the same energy input conditions, the several weight vectors were used to compare the far-field sound field control effects of low-frequency phased array. From the simulating results and the mean square deviation of time-delay estimation exited by the weight vectors, it was showed and proved that the phased array emitting the No.1 normal mode was not always the optimum in the far-field sound field control.

5a FRI. AM

5aUWa11. Spatial diversity of ambient noise in the new Arctic. Henrik Schmidt, Scott Carper, Thomas Howe (Mech. Eng., Massachusetts Inst. of Technol., 77 Massachusetts Ave., Rm. 5-204, Cambridge, MA 02139, henrik@mit.edu), and Andrew Poulsen (Appl. Physical Sci., Lexington, MA)

The Arctic Ocean is undergoing dramatic changes, the most apparent being the rapidly reducing extent and thickness of the summer ice cover. As has been well established over prior decades, the environmental acoustics of the ice-covered Arctic is dominated by two major effects: the highly inhomogeneous ice cover, and the monotonically upward refracting sound speed profile, the combination of which forces all sound paths to be exposed to strong scattering loss and the associated loss of coherence. In some portions of the Arctic Ocean, however, inflow of warm Pacific water has created the so-called "Beaufort Lens," a neutrally buoyant high sound velocity layer at 70-80 meter depth, which has dramatically altered the acoustic environment, creating a strong acoustic duct between approximately 100 and 300 m depth. This duct has the potential of trapping sound out to significant ranges (80-100 km) without interacting with the ice cover, resulting in much higher coherence and signal preservation. Acoustic noise measurement results collected with a vertically suspended array during ICEX 2016 illustrate the spatial and temporal noise properties in the presence of this acoustic duct at different depths. Comparisons of the ICEX 2016 data are also made with modeled Arctic noise data. [Work supported by ONR and DARPA.]

11:00

5aUWa12. Arctic ambient noise measurement comparisons in the Beaufort Sea. Andrew Poulsen (Appl. Physical Sci., 49 Waltham St., Lexington, MA 02421, poulsen@alum.mit.edu) and Henrik Schmidt (Massachusetts Inst. of Technol., Cambridge, MA)

During ICEX 2016, an MIT autonomous underwater vehicle (AUV) was deployed in the Beaufort Sea region of the Arctic Ocean with the technical objectives of demonstrating the deployment, operation and recovery of an AUV with a towed array under extreme under-ice Arctic conditions, and the scientific objective of characterizing the acoustic environment. Part of this effort included suspending the AUV from a hydro-hole with the acoustic array hanging in a vertical configuration. This new data set created a choice opportunity to reprocess similar vertical array data from the same region of the Beaufort Sea collected by MIT during SIMI 1994. In the intervening twenty-two years between these two ice camps, significant changes have occurred in the Arctic, including rapidly reducing extent and thickness of the summer ice cover. Furthermore, a persistent inflow of a shallow "tongue" of warm Pacific water has recently been discovered in the Beaufort Sea, creating a strong acoustic duct between approximately 100 and 200 m depth. This paper presents noise properties measured more than two decades apart in a region of rapid and significant change. [Work supported by ONR and DARPA.]

11:15

5aUWa13. Noise background levels and noise event tracking/characterization under the Arctic Ice Pack: Experiment, data analysis, and modeling. Kevin Williams, Michael L. Boyd, Alexander G. Soloway, Eric I. Thorsos, Steven G. Kargl, and Robert I. Odom (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, williams@apl.washington.edu)

In March of 2014 an Arctic Line Arrays System (ALAS) was deployed as part of an experiment in the Beaufort Sea. The background noise levels in the frequency range from 1 Hz to 25 kHz were measured. The goal was to have a 3d sparse array that would allow determination of the direction of sound sources out to 100s of km and both direction and range of sound sources out to 1-2 km from the center of the array. ALAS started recording data at 02:12 on March 10, 2014 (UTC). It recorded data almost continuously at a sample rate of 50 kHz until 11:04 on March 24, 2014. Background noise spectral levels are presented for low and high floe-drift conditions. Tracking/characterization results for ice cracking events, including the initiation of an open lead and a seismic event are presented. Results from simple modeling indicate that the signature of a lead formation may be a combination of both previously hypothesized physics and enhanced emissions near the ice plate critical frequency. The seismic T-wave arrival time indicates that a significant amount of energy coupled to T-wave energy somewhere along the path between the earthquake and ALAS.

11:30

5aUWa14. Environmentally adaptive autonomy for acoustic communication and navigation in the new Arctic. Henrik Schmidt (Mech. Eng., Massachusetts Inst. of Technol., 77 Massachusetts Ave., Rm 5-204, Cambridge, MA 02139, henrik@mit.edu) and Toby Schneider (Gobysoft LLD, Woods Hole, MA)

The particular sensitivity of the Arctic to climate change is well established. and the significance to undersea operations can be dramatic. As part of the recent ICEX16 US Navy Exercise in the Beaufort Sea, MIT deployed an autonomous underwater vehicle with a towed hydrophone array below the ice cover for assessing the climate induced changes to the undersea ambient noise environment. The safe underwater operation depended on navigation updates from the submarine tracking range being communicated to the vehicle for fusion with the onboard inertial navigation. However, the changes in the environment severely deteriorated the tracking performance compared to previous deployments. The reason was clearly associated with a previously observed neutrally buoyant layer of warm Pacific water persistently spreading throughout the Beaufort Sea, which severely alters the acoustic environment with dramatic effects for both long and short range acoustic sensing, communication and navigation. This paper describes the effects observed and discusses how robust acoustic connectivity in this environment makes it paramount that the manned or unmanned undersea platforms are capable of adapting to the environment for sensing, communication, and navigation. [Work supported by the Office of Naval Research.]

11:45

5aUWa15. Detection and communication in the "Beaufort Lens". Arthur B. Baggeroer, Henrik Schmidt, Scott A. Carper (Mech. and Elec. Eng., Massachusetts Inst. of Technol., Rm. 5-206, MIT, Cambridge, MA 02139, abb@boreas.mit.edu), Jon Collis, and Jordan Rosenthal (MIT Lincoln Lab., Lexington, MA)

The WHOI Ice Tethered Profilers in the Beaufort has confirmed a "sound speed duct" at depths between 100 to 250 m labeled as the "Beaufort Lens." It is thought to be caused by warm water intrusion from the Bering Strait. A significant consequence of the duct is the prediction of 10 dB lower transmission losses at ranges of 100 km for sources and receivers both within the duct. We examine how the properties of the duct impact detection and communication. Transmission loss is strongly affected by frequency. Ducted propagation is not supported below a modal cutoff and absorption losses become consequential at higher frequencies. For active sonar, where wideband waveforms are employed, dispersion effects and multipath become important. The shape of the duct, especially boundary gradients, impacts this dispersion. While a vertical line array spanning the duct can potentially exploit the multipath ray/mode coherence and the resolution of the array impacts whether coherent or incoherent combination should be employed to achieve desired gains. Data for validating and verifying long range transmission loss in the duct were acquired by two submarines participating in ICEX-16. This presentation examines the signal processing for detection and communication within the so called "Beaufort Lens."

Session 5aUWb

Underwater Acoustics: Ocean Ambient Noise: Sources, Variation, and Characterization II

D. Benjamin Reeder, Chair
Oceanography, Naval Postgraduate School, 833 Dyer Rd., Bldg. 232, SP-311B, Monterey, CA 93943

Contributed Papers

10:00

5aUWb1. Study for computation/measurement of underwater ship radiated noise toward marine environmental protection in Japan. Hikaru Kamiirisa, Nobuaki Sakamoto, Akiko Sakurada (Fluids Eng. & Hull Design Dept., National Maritime Res. Inst., Japan, 6-38-1 Shinkawa, Mitaka, Tokyo 181-0004, Japan, kamiirisa@nmri.go.jp), takahiro kijima (mlit, Tokyo, Japan), yoshihito tominaga (JSTRA, Tokyo, Japan), and naoya umeda (Osaka Univ., Osaka, Japan)

USRN has been a large interest among maritime industries in terms of mariners' comfort as well as environmental protection. This paper presents the characteristics of the measured USRN in full scale as well as the near field propeller cavitation noise estimated by viscous computational fluid dynamics (CFD) simulation. The USRN measurement was carried out according to the ISO standard protocol in deep water. The results show that the characteristics of USRN spectrum, differences of SPL in tonal noise depending on broadside and engine load are successfully captured for which the level of background noise is low enough to distinguish target noise and background noise. For the CFD simulation, two methodologies have been tested to predict frequencies and sound pressure level (SPL) of near field propeller cavitation noise, i.e., 1) direct estimation and 2) indirect estimation. The results are encouraging in that the present CFD simulation is capable of capturing sheet cavitation and resultant tonal noise. The estimation of broadband noise is still challenging by the present CFD simulation yet its application in conjunction with the semi-empirical formula is likely to be effective for predicting the upper bound level of broadband noise.

10:15

5aUWb2. The design and implementation of ocean ambient noise acquisition system based on the underwater glider. Lu Liu (Institute of Acoust.,Chinese Acad. of Sci., No. 21 North Sihuan West Rd., Haidian District, Beijing 100190, China, 13398625258@163.com)

Ocean ambient noise (OAN) is the background noise of the ocean,which is considered as the interference for the underwater acoustic equipment,on the other hand,it can be used to invert the parameters of the ocean environment. It is important to acquire and study the spatial temporal characteristics of ocean ambient noise. As a new type of underwater detection platform,underwater glider's performance is suitable to acquire the ocean ambient noise. This paper designed and realized an OAN acquisition system (OANAS),which is mounted on the Underwater glider. The system mainly includes two parts,the deepwater hydrophone array and the acquisition system. This design uses the ADI's BF518,the AD73360 and the uClinux to make the acquisition system have high speed,low power consumption and low noise abilities,it can also sample eighteen underwater signal roads synchronously and meet the requirement of the system. The data memorized is proved effective,the system is proved to be stable and meet the design requirement through the test which the OANAS mounted on the Hybrid Underwater Glider in the National Ocean Technology Center.

10:30

5aUWb3. Eigenvalues of the sample covariance matrix for a vertical array in shallow water. Jie Li and Peter Gerstoft (Marine Physical Lab., Scripps Inst. of Oceanogr., 9500 Gilman Dr. La Jolla, San Diego, CA 92093, jil004@ucsd.edu)

The eigenvalue spectrum of the sample covariance matrix (SCM) contains valuable information about the environment and the source and noise generation. Random Matrix Theory (RMT) combined with physical models is used to explain the eigenvalue distribution. For conventional three-dimensional and two-dimensional free space isotropic noise models, there is a jump in the eigenvalue spectrum of SCM which only depends on array properties (element spacing to wavelength ratio). This paper investigates the eigenvalue distribution of SCM for a vertical array in shallow water noise environment using Harrison's noise model [Harrison, J. Acoust. Soc. Am. **99**, 2055-2066 (1996)]. Results show that when bottom is considered, the jump location depends on both array and bottom properties. RMT is applied to study the eigenvalue spectrum estimation with different sound speed and attenuation in the bottom. Shallow water vertical array data are analyzed and it is demonstrated that the eigenvalues of the SCM compare well with theory.

10:45

5aUWb4. Investigating flow noise on underwater gliders acoustic data. Francisco A. dos Santos, Pedro M. São Thiago, André Luis S. de Oliveira, Rafael Barmak (PROOCEANO, Av. Rio Branco, 311/1205, Centro, Rio de Janeiro, RJ 20040-009, Brazil, francisco@prooceano.com.br), José Antônio M. Lima (CENPES, PETROBRAS, Rio de Janeiro, RJ, Brazil), Fernando G. de Almeida (UO-BS, PETROBRAS, Santos, SP, Brazil), and Thiago P. Paula (CENPES, PETROBRAS, Rio de Janeiro, RJ, Brazil)

Since November 2015, two underwater gliders equipped with external hydrophones were deployed in the South Brazilian Bight in order characterize the area's soundscape. Contrasting to standard fixed mooring systems (where flow noise is generated by currents passing by the hydrophones), gliders are subject to noise generated by its own downward and upward motion, which may compromise soundscape characterizations if not properly evaluated. In order to investigate induced flow noise on the hydrophone and its characteristics, 563 hours of acoustic data from the gliders were correlated to the navigation settings. Results can be comparable to previous flow noise descriptions for fixed systems. A high correlation was observed between the glider speed (both vertical and total) and 1/3 octave band levels centered at frequencies below 20 Hz. Estimates of the broadband sound pressure level were accomplished with different lower frequency limits and found to be uncorrelated to the glider speed above 40 Hz.

5a FRI. AM

5aUWb5. Acoustic data quality assessment tools and findings for ocean observing systems. Thomas Dakin, John Dorocicz, Ben Biffard (Ocean Networks Canada, TEF-128A 2300 McKenzie Ave, University of Victoria, Victoria, BC V8W2Y2, Canada, tdakin@uvic.ca), Nathan D. Merchant (Ctr. for Environment, Fisheries and Aquaculture Sci., Lowestoft, United Kingdom), David Hannay (JASCO Applied Sci., Victoria, BC, Canada), Steve Mihaly, Marlene Jeffries, and Jeannette Bedard (Ocean Networks Canada, Victoria, BC, Canada)

Ocean Networks Canada (ONC) operates long time series, ocean observatories in the Pacific and Arctic. These include the large VENUS and NEPTUNE observatories, many small community based observatories and the Underwater Listening Station (ULS) for the Vancouver Fraser Port Authority. Passive acoustic monitoring systems are a component of all ONC observatories and passive acoustic data quality is therefore a concern. All the observing systems have multiple underwater electronics and sensor types, many of which can negatively impact the passive acoustic sensor data. Hydrophone sensitivity degradation due to time, water absorption, and biofouling need to be assessed to ensure accurate ambient noise measurements and accurate vessel underwater radiated noise level measurements. The performance and suitability of the hydrophones for specific areas also needs to be assessed so the acoustic analysts can be aware of the hydrophone induced data limitations. ONC has been examining the use of in situ calibration verifications, spectral probability density (SPD) plots, spectrograms, and wave data as tools to assess the passive acoustic data quality. The preliminary findings on the impact of all of the above acoustic error sources are presented.

5aUWb6. Some new observations on the underwater sound field from impact pile driving. Peter H. Dahl and David R. Dall'Osto (Appl. Phys. Lab. and Mech. Eng., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, dahl@apl.washington.edu)

The underwater sound field from impact pile driving has been studied in recent years, motivated largely by the high levels and impulsive character of the sound and potential impacts on marine life. Here we report on new observations made on a vertical line array (VLA) and neutrally buoyant geophone (vector sensor) both positioned 120 m from the pile, to inform future modeling and mitigation efforts. These observations include: (1) a ground-wave precursor of amplitude 200 Pa, or 10% of peak pressure (2) delayed arrival structure in the impulsive time series due to injection of airborne sound from the hammering process, and (3) properties of the Scholte wave as a function of receiver depth. The study originates from Puget Sound, WA, and involves a 30-cm diameter steel pile installed with impact hammer energy ~200 kJ. Water depth at pile was 9 m, gradually increasing to 17 m at the VLA range. Pile depth was recorded over the sequence of 300 pile strikes, and this knowledge is essential to interpret these observations as the pile was driven 12 m into the sediment. [Research supported by Washington State Dept. of Transportation with partial support from ONR.]

5aUWb7. Optical detection of mach shock wave cones in water using refracto-vibrometry. Matthew T. Huber (Dept. of Phys., Rhodes College, 2000 North Parkway, Rhodes Box 2644, Memphis, TN 38112, hubmt-18@rhodes.edu), Nathan Huber, and Thomas Huber (Dept. of Phys., Gustavus Adolphus College, St. Peter, MN)

Recently, a hydrophone array was used to demonstrate Mach shock wave cone formation in water when 1m diameter steel pile rods for civil structures were driven into the ground by massive impact hammers. These shock wavefronts are of concern for marine mammals and fish. In the current project, Mach shockwave cones were directly imaged using refracto-vibrometry. A Polytec PSV-400 laser Doppler vibrometer was directed through a water tank towards a stationary retroreflective surface. The density variations of acoustic wavefronts which pass through the laser cause variations in the optical path length between the laser and retroreflector. This results in a time-varying modulation of the laser signal returning to the vibrometer, enabling optical detection of the acoustic wavefronts. A 35 mm diameter rod (steel or other metal) immersed in a water tank was repeatedly "impacted" by narrow pulses from a 1 MHz ultrasound transducer. The vibrometer sampled numerous scan points to generate videos of the time evolution of a Mach cone wavefronts launched as the compressional and shear waves travel through the rod. The angle of the Mach cone produced is consistent with the speed of waves in the rod.

5aUWb8. Statistical wavelet filtering for impulsive noise mitigation. Jit Sarkar, Christopher M. Verlinden, Jeffery D. Tippmann, William S. Hodgkiss, and William A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., 9500 Gilman Dr., Mail Code 0238, La Jolla, CA 92093-0238, bsarkar@ucsd.edu)

Impulsive noise sources in acoustic data can represent a significant amount of contamination that is not easily addressed by spectral methods alone. In ocean acoustics, such noise sources can be biological in nature, e.g., marine mammal localization clicks, or even mechanical noise induced on an instrument itself. While the broadband character of an impulse does not lend itself to frequency domain filtering, the wavelet domain is more appropriate and has been used to analyze impulsive sources of interest in the past. We present a method for filtering unknown nuisance impulses using basic statistics of the continuous wavelet transform (CWT), with minimal disturbance to the remaining signal(s) of interest. Examples from both simulation and collected ocean acoustic data are presented.

Session 5pABa

Animal Bioacoustics: Behavioral Response Studies

Rebecca Dunlop, Cochair
School of Veterinary Science, University of Queensland, Cetacean Ecology and Acoustics Lab, Gatton QLD 4343, Australia

Kotaro Ichikawa, Cochair
Field Science Education and Research Center, Kyoto University, Kitashirakawa Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan

Chair's Introduction—1:00

Invited Papers

1:05

5pABa1. Southern California behavioral response study—Overview and synthesis of results to date. Brandon L. Southall (SEA, Inc., 9099 Soquel Dr., Ste. 8, Aptos, CA 95003, Brandon.Southall@sea-inc.net), John Calambokidis (Cascada Res., Olympia, WA), Ari Friedlaender (Oregon State Univ., Newport, OR), Stacy DeRuiter (Calvin College, Grand Rapids, MI), Alison Stimpert (Moss Landing Marine Lab., Moss Landing, CA), Jeremy Goldbogen (Hopkins Marine Station, Stanford Univ., Monterey, CA), Elliott Hazen (UC Santa Cruz, Santa Cruz, CA), David Moretti (Naval Undersea Warfare Ctr., Newport, RI), Ann Allen, Glenn Gailey, Annie Douglas, Greg Schorr, Erin Falcone (Cascada Res., Olympia, WA), Dave Cade (Hopkins Marine Station, Stanford Univ., Monterey, CA), Fleur Visser (Kelp Marine Res., Amsterdam, Netherlands), and Jay Barlow (NMFS Southwest Fisheries Sci. Ctr., NOAA, LaJolla, CA)

The Southern California Behavioral Response Study (SOCAL-BRS) is an interdisciplinary, multi-team collaboration that uses high-resolution, multi-sensor tags to document behavioral responses of cetaceans to Navy mid-frequency (2.5-5 kHz) active sonar (MFAS). Individual animals are monitored before, during, and after controlled exposure experiments (CEEs) using either simulated or [the first-ever use of] actual U.S. Navy ship-based MFAS. Over 175 tags have been deployed on individuals of ten cetacean species and over 80 CEEs have been conducted. Results including baseline behavioral data, as well as behavioral responses to sound exposure have been published in more than a dozen scientific publications. Behavioral responses to MFAS and other mid-frequency sounds have been documented in several species with changes in diving and feeding behavior and avoidance of sound sources being the most common responses observed, though some individuals did not respond despite relatively high received sound levels. Responses appear to depend on species, exposure context (e.g., behavioral state, prey distribution, proximity to sound sources), and source type (simulated versus real). These experiments provide controlled, scientific measurements of response, or lack of response, that are directly applicable to improving Navy environmental compliance assessments of behavioral response to MFAS.

1:25

5pABa2. Baleen whale calling behavior and response to anthropogenic sound. Alison K. Stimpert (Moss Landing Marine Labs., 8272 Moss Landing Rd., Moss Landing, CA 95039, alison.stimpert@gmail.com), Stacy L. DeRuiter (Calvin College, Grand Rapids, MI), Erin A. Falcone (Marecotel, Seabeck, WA), John Joseph, Tetyana Margolina (Naval Postgrad. School, Monterey, CA), David Moretti (Naval Undersea Warfare Ctr., Newport, RI), Selene Fregosi, Ari S. Friedlaender (Oregon State Univ., Newport, OR), John Calambokidis (Cascadia Res. Collective, Olympia, WA), Peter L. Tyack (Scottish Oceans Inst., Univ. of St. Andrews, St. Andrews, Scotland, United Kingdom), Jeremy A. Goldbogen (Dept. of Biology/Hopkins Marine Station, Stanford Univ., Pacific Grove, CA), and Brandon L. Southall (Southall Environ. Assoc., Aptos, CA)

Without a means of studying large whales in a controlled experimental environment, less is understood about their sound production mechanisms than is understood about those of smaller odontocetes. To describe call production behavior in fin whales, we used a recent technique that correlates fast-sampling accelerometer signals from tags with concurrently recorded acoustic signals to identify calls produced by the tagged animal. We tagged 18 fin whales as part of the Southern California Behavioral Response Study (SOCAL BRS), of which four were confirmed to be calling. We were then able to describe their kinematic and social behavior in relation to call production. Behaviors associated with elevated call rates included shallow maximum dive depths, little body movement, and negative pitch in body orientation, similar to some other calling baleen whale species. These are the first descriptions of body orientation and dive depths at which fin whales are most likely to call. We also describe calling responses (or lack thereof) from blue and fin whales exposed to simulated mid-frequency active sonar. The call behavior characterizations presented here will help with predicting calling behavior from surface behavior, informing interpretation of passive acoustic data, and further investigating effects of anthropogenic sound on baleen whales.

5p FRI. PM

5pABa3. The behavioral response of humpback whales to seismic air gun noise. Rebecca Dunlop, Michael J. Noad (School of Veterinary Sci., Univ. of Queensland, Cetacean Ecology and Acoust. Lab., Gatton, QLD QLD 4343, Australia, r.dunlop@uq.edu.au), Robert McCauley (Curtin Univ., Perth, WA, Australia), and Douglas Cato (Univ. of Sydney, Sydney, NSW, Australia)

Four major experiments have been conducted off Australia to quantify the behavioral response of migrating humpback whales to various seismic air gun arrays. The first, using a 20 in^3 air gun, was used to develop the analysis framework, which was then applied to later experiments. The following two experiments tested a 4-step "ramp-up" procedure (20, 60, 140, and 440 in^3 array sequence) and a 140 in^3 array. Both studies found a change in movement behavior in response to the air guns, where groups deviated more from their course and made less progress towards the source. There was no evidence of any change in surface behaviors, including respiration rates. The final experiment involved ramp-up to a 3130 in full commercial array. In response, whales decreased their dive time and displayed an elevated respiration rate, more so to the full array phase. In this phase, they also changed their surface behavior, in that breaching rates were elevated but tail and pectoral slapping behaviors were reduced. Consistent with previous studies, whales also changed their movement behavior, more so during ramp-up. Therefore the full array elicited a greater variety and magnitude of behavioral changes than observed with the smaller air gun arrays.

5pABa4. Vocal response of dugongs (*Dugong dugon*) to playbacks of conspecific calls suggest ranging function of chirps. Kotaro Ichikawa, Nobuaki Arai (Field Sci. Education and Res. Ctr., Kyoto Univ., Kitashirakawa Oiwake-cho, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan, ichikawa.kotaro.5r@kyoto-u.ac.jp), and Kongkiat Kittiwattanawong (Phuket Marine Biological Ctr., Phuket, Thailand)

Dugongs vocalize bird-like calls, such as chirps and trills but the functional definitions of the calls are yet to be clarified. A series of playback experiments was conducted in Thai waters to investigate their call-back behavior. The population was exposed to 4 different playback stimuli; a recorded wild dugong's chirp, a synthesized down-sweep sound having similar frequencies to the dugong chirp, a synthesized constant-frequency sound, and no sound as a control. Vocalizing dugongs were localized using an array of stereo-underwater-recording systems (AUSOMS-Ds). Total of 4068 calls were observed in reaction to the stimuli. Wild dugongs vocalized more frequently after the playback of dugong chirps (2.8 calls/min) than those of constant-frequency (0.55 calls/min) and control (0.2 calls/min). Ratio of the dugong chirps to all of the call type increased during the playback period. Dugongs were localized on 52 occasions within 25 m range from the playback source. Source level and duration of the chirps from wild dugongs responding to the playback showed positive correlation with distances between the caller and the playback speaker. These results suggest that dugongs can advertise their relative locations by exchanging chirps. Frequency-modulated chirps may have facilitated ranging between individuals.

5pABa5. Response to playback test in the captive Amazonian manatees (*Trichechus inunguis*) in Brazil. Mumi Kikuchi (Kyoto Univ., Wildlife Res. Ctr., 2-24 Tanaka-Sekiden-cho, Sakyo-ku, Kyoto 606-8203, Japan, mumikomo@gmail.com), Diogo de Souza, and Vera M. da Silva (National Inst. of Amazonian Res. (INPA), Aquatic Mammals Lab. (LMA), Manaus, Brazil)

Previous studies suggested that manatee calls were primarily for communication and not for navigational purposes. In this study, the vocal response of captive Amazonian manatees to playbacks of several acoustic stimuli was investigated. Experiments were conducted using nine captive Amazonian manatees at the National Institute of Amazonian Research (INPA), Brazil, in 2014. All manatees, except one, are orphan calves rescued from the illegal hunting or incidental catch. They were kept in the outdoor pool with a group of 2-3 individuals, which are considered to be related. Manatees were exposed to five different playback stimuli: a recorded vocal from a related manatee, a recorded vocal from an unrelated manatee, a synthesized constant frequency based on the fundamental frequency of a related manatee vocal, a synthesized sound which entirely different from manatee vocal, and silence. A total of 58 playback sessions was conducted and 22,590 calls were recorded. While manatees showed inter-individual variability in the response to the playback stimuli, they tended to produce more vocalization during the playback of the related manatee vocal. They also tended to vocalize more to the playback of the constant frequency sound. Some individual showed strong reactions; touching or staying near the speaker during the experiment.

Contributed Papers

5pABa6. Baleen whale responses to a high frequency active pinger: Implications for upper frequency hearing limits. E. E. Henderson (Navy Marine Mammal Program, SSC Pacific, 53560 Hull St., Bayside Bldg Rm. 205, San Diego, CA 92152-5001, elizabeth.e.henderson@navy.mil), Alison Stimpert (Moss Landing Marine Lab., Santa Cruz, CA), and Amanda Debich (Scripps Inst. of Oceanogr., San Diego, CA)

In order to test the possibility of using high frequency pinger tags to track baleen whales on Navy instrumented ranges, three blue (*Balaenoptera musculus*) and one humpback (*Megaptera novaeangliae*) whales were exposed to two high frequency pingers. The pinger frequencies were 37 and 45 kHz, with estimated received levels between 106 and 134 dB re 1 μPa. The whales were monitored prior to the exposure in order to establish their behavioral state and to acclimate them to the presence of the boat. Two of the blue whales were deep feeding, while the third blue whale and the humpback whale were traveling with intermittent bouts of possible foraging near the surface. Each exposure lasted approximately 30—40 minutes, and the whales were observed for an additional two to three surfacing intervals

post-exposure. None of the blue whales demonstrated any behavioral response. The humpback whale's response was inconclusive, as other foraging animals entered the area at the same time; however, it is likely there was no response. These data may begin to provide information on the upper frequency limits of baleen whale hearing, as these species have responded to Navy sonar-like sounds at lower frequencies but similar received levels and behavioral states.

5pABa7. New parametric acoustic alarm proves effective at alerting wild manatees of approaching boats. Edmund R. Gerstein (Psych., Florida Atlantic Univ., 777 Glades Rd., Boca Raton, FL 33486, gerstein2@aol.com) and Laura Gerstein (Levaithan Legacy Inc., Boca Raton, FL)

The efficacy of a bow mounted parametric acoustic alarm for alerting manatees of approaching boats was demonstrated with wild manatees in Florida. Two experimental conditions were tested; (1) boat approaches with

the alarm-on, and (2) approaches with the alarm-off. Synchronized underwater acoustic buoys and aerial video recorders documented manatee behavior and acoustic conditions prior to, during, and after controlled boat approaches. 95% of the manatees during alarm-on trials elicited overt changes in behavior (avoidance reactions) prior to the boat reaching the focal manatee. In contrast, only 5% of manatees during alarm-off trials elicited any change in behavior prior to the vessel having to veer off. Changes in behavior observed prior to the boat reaching focal manatees was significantly greater (F = 471.9, df = 1, p< 0.01) during alarm-on trials. The mean distance at which manatees reacted to boat approaches during alarm-on trials was also 20 m compared to only 6m for alarm-off trials (F= 218.4 df = 1, p< 0.01). Manatees responded when received acoustic levels exceeded 12 dB of their critical ratios (behavioral threshold). Counter-intuitive to manatee protection laws, slow boats can be very difficult for manatees to detect and locate. Shallow water attenuates the dominant lower frequencies of slower moving propellers. The directional alarm assures manatees will be able to detect and locate boats at levels & distances sufficient to avoid injury. [Funded by DOD Legacy Natural Resource Management Program, USFWS Permit MA063561-4.]

3:30

5pABa8. Time-frequency analysis of bird songs and neurobehavioral study of the "Catbird Dumetella carolinensis". Babak Badiey (Newark Charter School, 20 Mcintire Dr., Newark, DE 19711, babak.badiey@gmail.com)

There is a commonality between the acquisition of spoken language in human infants and the acquisition of bird songs at the behavioral, neural, genetic, and cognitive levels. Similarities between how birds learn their language and how humans learn is believed to be fundamentally related to the way the brain works. Although the effects of the environment on hearing sensations, and hence the behaviors seems intuitive, it is not yet understood. In this paper, we use common digital recording systems to study bird songs in their natural habitat. Then by utilizing the principals of signal processing and spectral analysis individual syllabi are identified and studied for various songs. With the knowledge of the environment from direct measurement and the change in background noise level, the behavior of birds can be studied and documented. This technique was applied to recordings of Dumetella carolinensis (Gray Catbird) in the North East United States. It is shown that in spite of the natural instinct, the catbirds do not easily change their behavior at the onset of loud transient industrial sound. This may be related to the fact that the birds have already been exposed and have learned to live in noisy cosmopolitan environments.

3:45

5pABa9. Effects of industrial tones on bowhead whale calling rates. Susanna B. Blackwell (Greeneridge Sci., Inc., 90 Arnold Pl., Ste. D, Santa Barbara, CA 93117, susanna@greeneridge.com), Christopher S. Nations, Mandy E. Kauffman (WEST, Inc., Cheyenne, WY), Aaron M. Thode (Scripps Inst. of Oceanogr., San Diego, CA), Robert G. Norman, and Katherine H. Kim (Greeneridge Sci., Inc., Santa Barbara, CA)

In summer 2012, exploratory drilling was performed by Shell at Sivulliq, a lease holding in the Beaufort Sea, located within the autumn migration corridor of bowhead whales. The drilling operation involved a number of vessels performing various activities, such as towing the drill rig, anchor handling, and drilling. We aimed to assess the effect of sounds from these activities on bowhead whale calling rates. Acoustic data were collected with six arrays of directional recorders (DASARs) deployed on the seafloor over ~7 weeks in Aug-Oct. Industrial sound was quantified with the use of indices, that measured the presence and amplitude of tones from machinery, or the presence of airgun pulses. For each 10-min period of data collected at each of the 40 recorders, the number of whale calls detected was matched with the "dose" of industrial sound received, and the relationship between calling rates and industrial sound was modeled using negative binomial regression. The analysis showed that with increasing tone levels, bowhead whale calling rates initially increased, peaked, and then decreased. This dual behavioral response is similar to that previously described for bowhead whales and airgun pulses. [Work supported by Shell Exploration and Production Company.]

4:00

5pABa10. Opportunistic behavioral-response studies of baleen whales in response to US Navy Sonar Training off Kauai, Hawaii. Stephen W. Martin, Cameron R. Martin, Brian Matsuyama (National Marine Mammal Foundation, 2240 Shelter Island Dr., Ste. 200, San Diego, CA 92106, steve.martin@nmmpfoundation.org), Elizabeth E. Henderson, and Tyler Helble (SPAWAR Systems Ctr. Pacific, San Diego, CA)

Opportunistic behavioral responses of baleen whales to disturbances from US Navy mid-frequency active sonar (MFAS) training at the Pacific Missile Range Facility, Kauai, Hawai'i, are being studied utilizing passive acoustic recordings. Automated passive acoustic detection, classification, localization, and tracking analyses of the data have shown a behavioral response in terms of a reduction, or cessation, of minke whale "boing" calling in response to US Navy training during the month of February 2011, 2012, and 2013 over a study area of 3,780 km2. The reduced calling is expressed as reduced minimum densities in the study area by utilizing acoustically localized individual whale counts. In February 2011, the density before sonar training was 3.64 whales while the density during sonar training was 0.69 whales (95% confidence intervals of 3.31-4.01 and 0.27-1.8 whales, respectively). Individual ship-whale encounters have been observed to show cessation of calling from ship approaches without MFAS activity as well as ship approaches with MFAS. Sound Pressure Levels and Cumulative Sound Exposure Levels animals are exposed to are being estimated for evaluation of dose-response relationships. Tracking individual whales allows investigation of kinematics coupled with acoustic call details to establish baseline behaviors for comparison with observations during US Navy training.

4:15

5pABa11. Mid-frequency active sonar and beaked whale acoustic activity in the Northern Mariana Islands. Anne Simonis, Bruce Thayre (Scripps Inst. of Oceanogr., 9500 Gilman Dr., La Jolla, CA 92037-0205, asimonis@ucsd.edu), Erin Oleson (Pacific Islands Fisheries Sci. Ctr., National Oceanic and Atmospheric Administration, Honolulu, HI), and Simone Baumann-Pickering (Scripps Inst. of Oceanogr., La Jolla, CA)

Mid-frequency active (MFA) sonar has been associated with multiple mass stranding events of beaked whales around the world. A recent increase in military training exercises in the Mariana Archipelago corresponds with the presence of MFA sonar in the surrounding waters. We provide a quantitative report on MFA sonar and beaked whale acoustic activity detected on two autonomous acoustic recording packages deployed near Saipan and Tinian from March 2010 through December 2013. There were no detections of MFA sonar at Saipan during the 5-month deployment in 2010. On August 21, 2011, MFA sonar was detected near Saipan concurrent with a stranding event involving two Cuvier's beaked whales (*Ziphius cavirostris*). After one observed day of MFA sonar activity in Saipan and Tinian in 2011, observations increased to 1 month of ongoing activity at Saipan and nearly 3 months ongoing activity at Tinian in 2012. In 2013, MFA sonar events were observed during one day at Saipan and zero days at Tinian. Received levels, sound exposure levels, and temporal descriptions of the MFA sonar events are reported along with detections of beaked whale acoustic activity. Here, we highlight the importance of ongoing passive acoustic monitoring, especially for species like beaked whales that are difficult to visually detect at sea.

4:30

5pABa12. Marine mammal passive acoustics applied to the monitoring of long-term trends in beaked whale abundance and to the derivation of a behavioral risk function for exposure to mid-frequency active sonar. David Moretti (NUWC, 71 Woodmark Way, Wakefield, RI 02879, morettidj@hotmail.com), Tiago Marques (Univ. of St. Andrews, Lisbon, Portugal), Len Thomas (Univ. of St. Andrews, Saint Andrews, United Kingdom), Stephanie Watwood, Nancy DiMarzio, Karin Dolan, Ronald Morrissey, Jessica Shaffer, Joao F. Monteiro, and Susan Jarvis (NUWC, Newport, RI)

Knowledge of Cuvier's (*Ziphius cavirostris*) and Blainville's (*Mesoplodon densirostris*) beaked whales' distinct dive and vocal behavior has allowed for the development of multiple methods of passive acoustic abundance and density estimation (Marques *et al*., 2009, Moretti *et al*., 2010).

5p FRI. PM

These methods are being applied to multiple years of data to estimate long-term trends in abundance for Blainville's beaked whales at the Atlantic Undersea Test and Evaluation Center (AUTEC) in the Bahamas and at the Pacific Missile Range Facility (PMRF) in Hawaii, and for Cuvier's beaked whales at the Southern California Offshore Range (SCORE). These passive acoustic beaked whale dive data were combined with sonar and Range ship-track data to derive a behavioral risk function for Blainville's beaked whales at AUTEC, and are being extended to Cuvier's beaked whales at SCORE and Blainville's beaked whales at PMRF. The behavioral risk function maps the probability of beaked whale dive disruption as a function of the receive level of mid-frequency active sonar. The risk function and long-term trend analysis will help inform environmental policy going forward.

4:45

5pABa13. The effects of wind turbine noise on male Greater Prairie-Chicken vocal output during the breeding season. Cara Whalen, Mary B. Brown (School of Natural Resources, Univ. of Nebraska, Lincoln, NE), JoAnn McGee (Developmental Auditory Physiol. Lab, Boys Town National Res. Hospital, 555 North 30th St., Omaha, NE 68131), Larkin Powell (School of Natural Resources, Univ. of Nebraska, Lincoln, NE), and Edward J. Walsh (Developmental Auditory Physiol. Lab, Boys Town National Res. Hospital, Omaha, NE, edward.walsh@boystown.org)

At a previous meeting of the ASA, we reported that the Greater Prairie-Chicken (*Tympanuchus cupido pinnatus*) is, in all probability, able to detect noise produced by wind turbine installations, a finding based on turbine noise output and auditory brainstem response threshold measurements (Walsh *et al.*, 2015). That investigation was conducted in concert with an effort to determine if noise generated by a wind turbine farm located near Ainsworth, Nebraska affects the vocal output of male prairie chickens occupying breeding grounds (i.e., leks) where they perform elaborate physical and vocal mating displays. Whalen *et al.* (2014) paved the way to address this question by characterizing the acoustic properties of booms, cackles, whoops and whines, the primary call types produced by male birds, at sites remote from the turbine farm. At lek sites located relatively close (<1 km) to the wind farm installation, boom and whoop production levels were higher, boom durations were shorter, whine fundamental frequency was higher, and cackle biphonation occurred less frequently than among birds occupying leks located at remote locations. Although differences in vocal characteristics were statistically significant, vocal adjustments were relatively minor and the significance of wind turbine farm noise on call acoustics will be discussed. [Funded in part by USFWS-WSFR Wildlife Restoration Project W-99-R administered by the NGPC.]

5:00–5:15 Panel Discussion

FRIDAY AFTERNOON, 2 DECEMBER 2016　　　　　　　　　　　　　NAUTILUS, 1:00 P.M. TO 4:30 P.M.

Session 5pABb

Animal Bioacoustics: The Diversity of Bioacoustic Signals

T. Aran Mooney, Chair
Biology Department, Woods Hole Oceanographic Institution, 266 Woods Hole Rd., Woods Hole, MA 02543

Contributed Papers

1:00

5pABb1. Prusten and the acoustic character of socializing tigers. Adam Smith, JoAnn McGee (Developmental Auditory Physiol. Lab, Boys Town National Res. Hospital, 555 North 30th St., Omaha, NE 68131), Douglas Armstrong (Henry Doorly Zoo and Aquarium, Omaha, NE), and Edward J. Walsh (Developmental Auditory Physiol. Lab, Boys Town National Res. Hospital, Omaha, NE, edward.walsh@boystown.org)

Although tigers are generally solitary animals, acoustic communication among conspecifics plays an important role in their biology and behavior. This investigation was undertaken to extend understanding of prusten, one of the most common vocalizations produced by *Panthera tigris*. Prusten is a low-level, social vocalization uttered in close proximity of conspecifics. Calls were recorded from a group of nine captive adult tigers, representing the Amur (*P. tigris altaica*), Bengal (*P. tigris tigris*) and Malayan subspecies (*P. tigris jacksoni*). As described previously, the acoustic character of prusten consists of three to nine brief rhythmic pulses emitted at a mean rate of 11.5 pulses per second with a total mean duration of 0.5 seconds. Calls were generally low level, with a mean of 71 dB SPL re 20 μPa at 1 meter from the source. Spectral energy was broadband, extending from below 20 Hz to above 22 kHz, although low frequency energy dominated the call with mean peak and centroid frequencies of 130 and 987 Hz, respectively. A small subset of calls contained infrasonic energy. This study affirms and extends findings from previously published descriptions and is the first report of the production of infrasound in this low level, pulsatile vocalization. [Work supported in part by NSF Grant Award #0823417.]

1:15

5pABb2. Using CT to predict vocal tract resonances in mammal vocalizations: Are calls nasalized? David Reby (School of Psych., Univ. of Sussex, University of Sussex, Brighton BN2 9TJ, United Kingdom, reby@sussex.ac.uk), Megan T. Wyman (Dept. of Evolutionary Biology and Environ. Studies, Univ. of Zurich, Zurich, Switzerland), Roland Frey (Leibniz Inst. for Zoo and Wildlife Res. (IZW), Berlin, Germany), and Joel Gilbert (Laboratoire d'Acoustique de l'Université du Maine – UMR CNRS, Université du Maine, Le Mans, France)

Males of several species of deer have a descended larynx, which gives them an unusually long vocal tract (VT). They can extend their VT by further lowering their larynx during the production of their sexual loudcalls. Formant frequencies are lowered as the vocal tract is extended, as predicted when approximating the vocal tract as a uniform quarter wavelength resonator. However, formant frequencies in polygynous deer follow uneven distribution patterns, suggesting that the vocal tract shape is in fact rather complex. We CT-scanned the artificially extended vocal tract of two adult fallow deer, and measured the cross-sectional area of the supralyngeal vocal

tract along the oral and nasal pathways. We used this data to model resonances patterns produced by these VT including the oral pathway, the nasal pathway, or both. We found that the combined oral/nasal VT produced a formant pattern that more closely matches that observed in fallow deer groans and enables a better estimation of VT length from formant patterns. This clear indication that the nasal cavity and oral cavity are both simultaneously involved in the vocal production of a nonhuman mammal suggest that the potential for partial nasalization of putative oral loud calls should be carefully considered.

1:30

5pABb3. Restoring dueting behavior in a mated pair of buffy cheeked gibbons after exposure to construction noise at a zoo through playbacks of their own sounds. Jeanette A. Thomas, Brett Friel, and Sarah Yegge (Biological Sci., Western Illinois Univ.-Quad Cities, WIUQC, 3300 River Dr., Moline, IL 61265, jeanettethomaswiu@gmail.com)

There is concern about anthropogenic noise effecting zoo animals. Gibbons are known for duets exhibited by a mated pair. During a 6-month period in 2009, Niabi Zoo conducted construction (sewer/water lines, and paving) within 100 m of the buffy cheeked gibbon exhibit. Prior to construction, a male/female pair (plus their 4-year old son) performed loud, elaborate duets each day. Friel collected behavior data using focal animal and instantaneous scan-sampling for 10 days; 5 times per day prior to construction and for 20 days during construction. SEL measurements were taken during both periods. In general, the immature gibbon vocalized little; the female vocalized most. Gibbons vocalized less during construction noise. After construction, the subadult male became ill and was euthanized; thereafter the pair ceased vocalizing. Two years later, Yegge played back sounds to encourage dueting: 1) their own duet, 2) the duet of wild gibbons, 3) rock music, and 4) a silent control. Their behavior was collected for 3, 15-min playback sessions per day for 30 days. The pair's own playback elicited a duet even during the first session. The female vocalized significantly more than the male, especially during the wild gibbon duet. Currently, the pair continues dueting on their own.

1:45

5pABb4. Baby's first words: Vocal behavior and ontogeny of Northern right whales in the southeast critical habitat. Edmund R. Gerstein (Psych., Florida Atlantic Univ., 777 Glades Rd., Boca Raton, FL 33486, gerstein2@aol.com), Vasilis Tyrgonis (Univ. of the Agean, Lesvos, Greece), and James Moir (Marine Resource Council of Florida, Palm Bay, FL)

North Atlantic right whales are one of the most endangered of the great whales. A remnant population of ~500 inhabits the eastern seaboard of North America. A small fraction (2%) travels south to their critical calving habitat along the Florida and Georgia coast. By late November and through the winter, right whales give birth and nurse their calves in these shallow waters before departing in early spring to their northern habitats. In the southeast critical habitat mother-calf pairs remain generally isolated from other whales, presenting a unique platform to study vocal development and learning in large whales. Using small boats, GPS-instrumented, free-drifting autonomous acoustic buoys were deployed in close proximity to 44 photo-identified mother-calf pairs over 7 calving seasons. Surface video and synchronized underwater recordings documented their social and vocal behavior. With the exception of some low-energy gunshot sounds, mothers, and their calves, remained predominantly silent during the first 4 weeks. This might be due to calf maturation, and or a strategy to avoid harassment by other whales or potential predators. Over 100 calls have been analyzed from 15 different calves. Some of these calves were resampled at different stages at <1 week up to 12 weeks of age. Evidence of age-related variance and changes in call structure, complexity, power, rates, as well as vocal mimicry and learning are presented. [Funding: HBOI Florida PFW License Plate Fund, The Harry Richter Foundation and IBM, NOAA Permit #14233.]

2:00

5pABb5. Dwarf sperm whale (*Kogia sima*) echolocation clicks from Guam (Western North Pacific Ocean). Karlina Merkens (CRP, NOAA/PIFCS (Lynker Tech.), 1845 Wasp Blvd., Bldg. 176, Honolulu, HI 96818, karlina.merkens@noaa.gov), Yvonne Barkley, Marie Hill (CRP, NOAA/PIFSC (JIMAR), Honolulu, HI), and Erin Oleson (CRP, NOAA/PIFSC, Honolulu, HI)

The cryptic species of the genus Kogia, including the dwarf sperm whale (*Kogia sima*) and the pygmy sperm whale (*Kogia breviceps*), are very difficult to observe in any but the most calm sea conditions. However, recordings of signals from wild and captive animals reveal that they echolocate at high frequencies (peak frequencies > 100 kHz) which makes passive acoustic monitoring (PAM) a possibility. We present details from a recent encounter with *K. sima* in the wild near the island of Guam (Western North Pacific Ocean). Three individuals were observed during a small-boat, visual survey in May 2016, and recordings were collected using a Compact Acoustic Recording Buoy (CARB). These clicks, with mean peak frequency of 126 kHz (+/- 4.3 kHz), mean click duration of 72 us (+/- 21 us), and -3 dB bandwidth 5.5 kHz (+/- 1.6 kHz), had similar properties to recordings of wild *K. sima* from the Bahamas (Atlantic Ocean), and also appear similar to published recordings from *K. breviceps* (Madsen *et al.* 2005). Available data and analyses to date do not allow for absolute determination between the two Kogia species at this time, but recordings such as these bring us closer to definitive species classification.

2:15

5pABb6. Inter and intra specific variation in echolocation signals among cetacean species in Hawaii, the northwest Atlantic and the temperate specific. Tina M. Yack, Kerry Dunleavy, Julie N. Oswald (Bio-Waves, Inc., 364 2nd St., Ste. #3, Encinitas, CA 92024, tina.yack@bio-waves.net), and Danielle Cholewiak (NOAA Northeast Fisheries Sci. Ctr., Woods Hole, MA)

Odontocete species use echolocation signals (clicks) to forage and navigate. The aim of this study is to explore inter- and intra-specific variation in clicks among odontocete species in the Northwest Atlantic, Temperate Pacific, and Hawaii. Clicks were examined for seven species of delphinids in the western North Atlantic; common dolphin, Risso's dolphin, pilot whale, rough-toothed dolphin, striped dolphin, Atlantic spotted dolphin, and bottlenose dolphin. Newly developed PAMGuard tools were used to automatically measure a suite of click parameters. Five parameters were compared between species; duration, center frequency, peak frequency, sweep rate, and number of zero crossings. Significant differences in duration, center and peak frequency were evident between species within regions (Dunn's test with Bonferroni adjustment $p<0.05$). Geographic variation in click parameters between the three study regions was compared for five species; bottlenose dolphin, common dolphin, striped dolphin, pilot whale, and Cuvier's beaked whale. Significant differences in several parameters were found for all species between the regions (Dunn's test with Bonferroni adjustment $p<0.05$). These results suggest that there are species specific differences in clicks among delphinids and that geographic variation exists for multiple species. The ecological significance of these findings will be discussed along with implications for classifier development.

2:30–2:45 Break

2:45

5pABb7. Determining killer whale (*Orcinus orca*) call variability from passive acoustic monitoring in the Chukchi and Bering Sea, Alaska. Brijonnay C. Madrigal (Vertebrate Ecology Lab., Moss Landing Marine Lab., 8272 Moss Landing Rd., Moss Landing, CA 95039, bmadrigal@mlml.calstate.edu), Catherine L. Berchok, Jessica L. Crance (Marine Mammal Lab., Alaska Fisheries Sci. Center/NMFS/NOAA, Seattle, WA), and Alison K. Stimpert (Vertebrate Ecology Lab., Moss Landing Marine Lab, Moss Landing, CA)

Killer whales (*Orcinus orca*) are a highly vocal species that produce three types of vocalizations; pulsed calls, whistles, and clicks. Unlike the Northern and Southern Resident populations of western Canada and the Pacific Northwest, little is known regarding the acoustic behavior of resident

and transient killer whale populations north of the Aleutian Islands in the Bering and Chukchi Seas. Acoustic data were analyzed from moored recorders deployed by the Marine Mammal Laboratory at two sites each in the Bering and Chukchi Seas (BOEM-funded). The recorders sampled at 4 kHz on a 7% duty cycle (Bering) or 16 kHz on a 28% duty cycle (Chukchi). Over 1100 calls were identified, and discrete call classification was conducted using an alphanumerical system that distinguished calls by location, general contour, and segment variation. Parameters analyzed included call duration, start/end frequency, and delta frequency/time; periods of call repetition were common. These results will help determine if resident populations occur in the Chukchi Sea, and identify which transient populations are present. This initial work classifying killer whale sounds in the Arctic and along the Bering Sea shelf will also facilitate comparisons of call types within and among transient and resident killer whale populations.

3:00

5pABb8. Whistle characteristics of newly defined species, the Burrunan dolphin (*Tursiops australis*), in coastal Victorian waters in Australia. Rebecca Wellard (Ctr. for Marine Sci. and Technol., Curtin Univ., GPO Box U1987, Perth, WA 6845, Australia, becwellard@gmail.com), Kate Charlton-Robb (Australian Marine Mammal Conservation Foundation, Hampton East, VIC, Australia), Christine Erbe (Ctr. for Marine Sci. and Technol., Curtin Univ., Perth, WA, Australia), and Bob Wong (School of Biological Sci., Monash Univ., Clayton, VIC, Australia)

A newly defined species, the Burrunan dolphin (*Tursiops australis*), was described in 2011 by Charlton-Robb *et al.*, and is endemic to southern and south-eastern Australian coastal waters. This species' distribution is characterized by small, isolated, and genetically distinct populations. With only two known populations in Victoria, the species is now listed as "Threatened" under the *Victorian Flora and Fauna Guarantee Act*. Describing and quantifying the vocal repertoire of a species is critical for subsequent analysis of signal functionality, geographic variation, social relevance, and identifying threats associated with anthropogenic noise. Here, we present the first quantitative analysis of whistle characteristics for the species, undertaken on these endemic Victorian populations. Vocalizations of *T. australis* were recorded during population based surveys in 2007 and 2014 across the Gippsland Lakes and Port Phillip Bay, Victoria. Acoustic parameters of whistles were measured including minimum frequency (fmin), maximum frequency (fmax), start frequency (sf), end frequency (ef), delta frequency (df= fmax-fmin), duration, number of extrema, number of inflection points, and number of steps. We review and compare *T. australis* whistle features to the two other bottlenose dolphins, *T. truncatus* and *T. aduncus*, to assess the similarity and/or differences between the sounds of the three species of bottlenose dolphins.

3:15

5pABb9. Visual acoustic analysis techniques for North Atlantic right whale sounds. James A. Theriault, Gary Inglis (Underwater Sensing, Defence Res. and Development Canada - Atlantic Res. Ctr., 1 Challenger Dr., Dartmouth, NS B2Y 4A2, Canada, jim@theriault-family.ca), and Hilary B. Moors-Murphy (Ocean and Ecosystem Sci. Div., Maritimes Region Fisheries and Oceans Canada, Dartmouth, NS, Canada)

Passive acoustic monitoring (PAM) is becoming a more widely accepted tool in mitigating the potential impact of man-made noise on marine mammals. Many marine mammals, and in particular cetaceans (whales and dolphins), use sound to communicate, navigate, forage, and avoid predators. Automated vocalization detectors have been in development for many years. Validation is achieved through aural/visual verification by analysts examining acoustic data in both the time and frequency domains. However, many factors also influence the ability of the analyst to properly detect and classify marine mammal sounds. First, a priori knowledge of the acoustic signature and the analysts experience are critical factors. Assuming the a priori knowledge exists, the ability to correctly assess the presence of call is dependent on the signal processing and displays available. Frequency resolution, temporal integration, and normalization either enhance or inhibit the ability to make the assessment. The signal processing parameterization varies between species and vocalization types. Six types of vocalizations have been previously described for the North Atlantic right whale (*Eubalaena glacialis*); upcalls, gunshots, screams, downcalls, blows, and warbles. The

parameterization required to optimally assess each of the vocalization types will be examined.

3:30

5pABb10. Song learning in humpback whales: Lessons from song hybridization events during revolutionary song change. Ellen C. Garland, Luke Rendell (School of Biology, Univ. of St. Andrews, St. Andrews, Fife KY16 9TH, United Kingdom, ecg5@st-andrews.ac.uk), M. Michael Poole (Marine Mammal Res. Program, Maharepa, French Polynesia), and Michael J. Noad (Cetacean Ecology and Acoust. Lab., School of Veterinary Sci., The Univ. of Queensland, Gatton, QLD, Australia)

Humpback whale songs are one of the most startling examples of transmission of a cultural trait and social learning in any non-human animal. Here, we investigate extremely rare cases of song hybridization, where parts of an existing song are spliced with a novel, revolutionary song, to understand how songs are learnt. Song unit sequences were extracted from over 800 phrases recorded during a song revolution (French Polynesia 2005), to allow fine-scale analysis of composition and sequencing. Clustering of song sequences (i.e., phrases) using the Levenshtein distance indicated songs clustered into three song types; a single hybrid phrase was identified representing the transition of one singer between two of these song types. A predictive model was fitted to the data and tested against the only other known recordings of humpback song hybridization: the eastern Australia 1996-97 song revolution. Songs change during revolutions through combining multiple complete phrases and themes from one song type, before transitioning through a hybrid phrase into the phrases and themes of a second song type. These extremely rare snapshots of song change represent the only known examples of song learning in over 20 years of data, spanning five populations of humpback whales were song revolutions occur.

3:45

5pABb11. Creation of an acoustic dictionary to classify the vocal repertoire of humpback whale song. Jennifer Allen (School of Veterinary Sci., Univ. of Queensland, Bldg. 8106, University of Queensland Gatton Campus, Gatton, QLD 4343, Australia, j.allen3@uq.edu.au), Ellen Garland (Univ. of St. Andrews, St. Andrews, United Kingdom), Anita Murray (School of Veterinary Sci., Univ. of Queensland, Brisbane, QLD, Australia), Michael Noad, and Rebecca Dunlop (School of Veterinary Sci., Univ. of Queensland, Gatton, QLD, Australia)

Animal vocal repertoires can be varied and complex, and it is often difficult to obtain a comprehensive picture. One of the primary reasons for this difficulty stems from the classification of acoustic signals, which has typically relied on qualitative and subjective methods. Vocal signals (units) were measured from east Australian humpback whale song from 2002-2014 and classified using self-organising maps (SOM). Humpback song was used here to provide an example of a particularly complex display of vocal signals. The use of SOM methodology takes an important step towards a more objective classification system. The resulting classifications were used to create a set of 149 hypothetical "exemplar" units to serve as an acoustic dictionary, representing the song's repertoire. The intent is to use this dictionary to determine vocal repertoire, as well as to facilitate the transcription and sequential analysis of large numbers of recordings with only a small representative subset of units requiring intensive acoustic feature measurement. Acoustic dictionaries are integral to studies of information content, microstructure, and syntax, and this methodology is thus applicable across the vocalization research community.

4:00

5pABb12. Acoustic characteristics of humpback whale (*Megaptera novaeangliae*) mother and calf vocalizations in the Hawaiian wintering grounds. Jessica Chen (Hawaii Inst. of Marine Biology, Univ. of Hawaii at Manoa, 46-007 Lilipuna Rd., Kaneohe, HI 96744, jchen2@hawaii.edu), Adam A. Pack (Departments of Psych. and Biology, Univ. of Hawaii at Hilo, Hilo, HI), Alison K. Stimpert (Moss Landing Marine Labs., Moss Landing, CA), and Whitlow W. Au (Hawaii Inst. of Marine Biology, Univ. of Hawaii at Manoa, Kaneohe, HI)

Humpback whale calves in the winter breeding grounds vocalize and the spectral characteristics of some of these vocalizations have been described,

but sound levels of calf vocalizations have not been investigated. There is also a lack of general information on vocalizations from mothers. To address these issues, we deployed suction cup acoustic and movement recording tags on humpback whale calves, mothers and a lone post-yearling stage female on the Hawaiian breeding grounds to record vocalizations and percussive sounds. Deployments took place in waters off West Maui over four winter seasons. Tags were deployed on 7 humpback whale mothers, 6 calves, and one lone female for approximately 57 hours of recordings. Calling rates of tagged animals were relatively low compared to song, with individual's means ranging from 0 to 16 vocalizations per hour. However, most calls occurred singly or in bouts, with long periods of silence before the next vocalization or set of vocalizations. These included sounds resembling non-calf social sounds as well as single song units, with durations up to 2 seconds and fundamental frequencies below 1.5 kHz. Our findings provide information that is critical to understanding vocal development in humpback whale calves and the sounds produced by adult females.

4:15

5pABb13. Loud and clear: Characterization of particle motion in Humpback whale song and its potential role in communication. T. Aran Mooney, Maxwell B. Kaplan (Biology Dept., Woods Hole Oceanographic Inst., 266 Woods Hole Rd., Woods Hole, MA 02543, amooney@whoi.edu), and Marc O. Lammers (Hawaii Inst. of Marine Biology, Univ. of Hawaii, Kaneoho, HI)

Acoustic communication can rapidly transfer a substantial amount of information, yet emitted signals must be conveyed with enough clarity to allow appropriate responses. Many mysticete calls such as humpback whale (*Megaptera novaeangliae*) song can be detected over large distances as a result of the propagating acoustic pressure wave, yet little is known regarding the particle motion component of these signals. To explore the particle velocity of humpback whale song, three singing whales were recorded from a vessel off Maui, HI in March 2015, using a sensor that contained a digital magnetometer, tri-axial accelerometer, and an omnidirectional hydrophone. The median magnitude of the acoustic particle velocity signal was substantial (64.5 dB re: 1m/s) for song components with a median pressure of 135.4 dB re: 1μPa. As vessel and sensor gradually drifted away or toward the whales, acoustic particle velocity and sound pressure correspondingly decreased or increased (range: 49.3-77.9 dB re: 1m/s and 118.4-148.1 dB re: 1μPa). The particle velocity signals were high even at 200m from the whale indicating this cue is substantial well into the far-field. There were predictable trends in the acoustic particle motion signal component and this component of song is a cue available to nearby animals.

FRIDAY AFTERNOON, 2 DECEMBER 2016   CORAL 1, 1:00 P.M. TO 5:15 P.M.

Session 5pBA

Biomedical Acoustics, Engineering Acoustics and Signal Processing in Acoustics: Acoustic Imaging

Yasuhiro Oikawa, Cochair
Department of Intermedia Art and Science, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan

Hideyuki Hasegawa, Cochair
University of Toyama, 3190 Gofuku, Toyama 9308555, Japan

Invited Papers

1:00

5pBA1. Signal processing for optical sound field measurement and visualization. Kohei Yatabe, Kenji Ishikawa, and Yasuhiro Oikawa (Intermedia Art and Sci., Waseda Univ., 59-407 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan, k.yatabe@asagi.waseda.jp)

Accurately measuring sound pressure is not an easy task because every microphone has its own mechanical and electrical characteristics. Moreover, the existence of a measuring instrument inside the field causes reflection and diffraction which deform the wavefront of sound to be measured. Ideally, a sensing device should not have any characteristic nor exist inside a measuring region. Although it may sound unrealistic, optical measurement methods are able to realize such ideal situation. Optical devices can be placed outside the sound field, and some of the sensing techniques, which decode information of sound from the phase of light, are able to cancel optical and electrical characteristics. Thus, optical sound measurement methods have possibility of achieving higher accuracy than ordinary sound measurement in principle. However, they have two main drawbacks that have prevented their applications in acoustics: (1) pointwise information cannot be obtained directly because observed signal is spatially integrated along the optical path; and (2) increasing signal-to-noise ratio is difficult because optical measurement of less than a nanometer order is typically required. To overcome the above difficulties, we have proposed several signal processing methods. Here, those methods are introduced with the physical principle of optical sound measurement.

5pBA2. Position and velocity measurement for moving objects by pulse compression using M-sequence-coded ultrasound. Shinnosuke Hirata and Hiroyuki Hachiya (Dept. of Systems and Control Eng., Tokyo Inst. of Technol., 2-12-1 Ookayama, S5-17, Meguro, Tokyo 152-8550, Japan, shin@ctrl.titech.ac.jp)

The pulse-echo method is one of the typical methods for ultrasonic distance measurement. Pulse compression using an M-sequence can improve the SNR of the echo reflected from an object and distance resolution in the pulse-echo method by the cross correlation between the received signal and the reference signal which correspond to the transmitted M-sequence-coded signal. In the case of a moving object, however, the echo is modulated due to the Doppler effect. The Doppler-shifted M-sequence-coded signal cannot be correlated with the original reference signal. Therefore, Doppler velocity estimation by autocorrelation of cyclic M-sequence-coded signal and cross correlations with Doppler-shifted reference signals which correspond to estimated Doppler velocities has been proposed. In this paper, measurement of position and velocity of the static and moving objects by the proposed method using the loud speaker and the microphone array are described. First, the Doppler velocity of the moving object is estimated in each array element. Second, received signals are correlated with the original reference signal for the static object. Then, received signals are correlated with Doppler-shifted reference signals for the moving object. Finally, positions of both objects are determined from the B-mode image formed by the synthetic aperture focusing technique.

5pBA3. A trial on the measurement of the acoustic properties by using a parametric loudspeaker. Akiko Sugahara (Eng., the Univ. of Tokyo, 4-6-1, Komaba, Meguro, Tokyo 153-8505, Japan, sugahara@iis.u-tokyo.ac.jp), Hyojin Lee, Shinichi Sakamoto (Inst. of Industrial Sci., the Univ. of Tokyo, Meguro, Tokyo, Japan), and Shigeto Takeoka (Shizuoka Inst. of Sci. and Technol., Fukuroi-shi, Shizuoka, Japan)

Measurement of acoustic properties of materials, such as absorption and reflection coefficient in a free field, has the results largely affected by diffraction by the sample edge. In order to minimize this, the use of a parametric loudspeaker, which enables to produce a directional audible sound by using the nonlinear effect of sound beams based on the parametric array theory, is investigated. At first, the directivity and the distance attenuation were measured as fundamental characteristics of the sound source. Then, the directivity of reflected sound from a surface, flat and diffuser plate, under oblique incidence conditions was measured in an anechoic chamber. The data were compared with those obtained by an experiment using a conventional loudspeaker, and with a wave-based numerical analysis using the mode matching method. The results show that the narrow sound directivity of the parametric loudspeaker is capable of obtaining the directivity properties of reflection sound from the plates more accurately than a conventional loudspeaker.

Contributed Papers

5pBA4. Non-contact acoustic imaging using single element transducer applied by time reversal focusing. Hideyuki Nomura and Hirotaka Saitoh (The Univ. of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi 182-8585, Japan, h.nomura@uec.ac.jp)

A large number of transducers have been used for the Time reversal (TR) processing to focus sound at a specific position and a specific time. In this study, TR focusing of ultrasound in space and time domains was experimentally confirmed in air using sound waves radiated from single element transducer. Experiments were carried out using a piezoelectric transducer glued on inner wall of a cavity for the reverberations of sounds. A microphone indicated that time reversed ultrasounds at about 50 kHz radiated from the hole of the cavity were focused at a specific time and a specific position outside of the cavity. This proposed TR focusing was applied to acoustic imaging for a defect on a brass plate with 0.3 mm in thickness. Time reversed ultrasound was focused at arbitrary position on the plate, and the laser Doppler vibrometry measured the vibration of the plate. Two-dimensional distribution of the vibration obtained by TR focusing indicated the presence of added small brass piece on the plate as a defect. These results suggest that the feasibility of single element TR focusing of ultrasound and an application of that on non-contact acoustic imaging.

5pBA5. Design of a low profile conformal array for transcranial ultrasound imaging. Aref Smiley, Mark Howell, Aaron Fleischman, Qi Wang, and Gregory T. Clement (Biomedical Eng., Cleveland Clinic, 9500 Euclid Ave., ND20, Cleveland, OH 44195, aref.smiley@gmail.com)

In our ongoing work toward transskull ultrasound brain imaging, we investigate an optically-registered "belt-type" conformal transducer designed to maximize energy transfer while maintaining precise phase control. Previous simulations determined that approximately 500 channels is sufficient for two-dimensional brain imaging by utilizing a tomographic approach valid on irregular boundaries [*Clement,* Inverse Problems*30 125010 (2014)*]. The present design is based around an assembly of rigid planar sub-arrays with flexible interconnects. Each piezo-composite sub-array (500 kHz, 1-3 random-fiber, 20 mm X 20 mm) is formed by electroding both faces of the composite before partially dicing the rear surface to produce 19 signal electrode terminals (width~0.85 mm, height~20 mm, and kerf~0.18 mm) with a shared ground face. Sub-arrays are bonded to a custom printed circuit boards and wired to shielded ribbon cable. The full transducer is then created by UV bonding the electrodes to a flexible band, forming the front transducer surface. Length may be customized ~500 elements for an average-sized head. Element performance is assessed by underwater scanned hydrophone characterization. Measurements numerically back-projected to the transducer face indicate good element isolation. Signals through an anatomically-correct full head phantom will be presented as well as the methodology for registering the array to the head using a 3D optical scanner.

5pBA6. Simulation of shear wave elastography imaging using the toolbox "k-wave". Fabrice Prieur (Dept. of Informatics, Univ. of Oslo, P.O Box 1080, Blindern, OSLO 0316, Norway, fabrice@ifi.uio.no), Bradley Treeby (Dept. of Medical Phys. and BioEng., Univ. College London, LONDON, United Kingdom), and Stefan Catheline (INSERM, LabTau, U1032, LYON, France)

Elastography is used to map the local elasticity of tissue. It can detect areas inside the body with a different elasticity from that of surrounding tissue indicative of pathologies like tumors. Shear wave elastography imaging produces an elasticity map by analyzing the propagation of shear waves. One source of shear wave is the acoustic radiation force produced by ultrasound. The prediction of the acoustic radiation force and of the shape and amplitude of the ensuing shear displacement is crucial. In this study, we present simulations of the radiation force produced by an ultrasound transducer and of the shear displacement it produces using the software package "k-wave." Results from simulations in a homogeneous and isotropic medium are compared against analytical solutions and results from a finite element modeling software. The obtained shear displacements from k-wave are very similar to the analytical solution with a root mean square error around

the focal zone below 9% and 21% for short and large propagation times, respectively. K-wave appears to be an accurate and efficient tool for simulation of acoustic radiation force and shear wave propagation. It combines the simplicity of finite time differences methods with the flexibility to simulate in any heterogeneous medium.

2:45

5pBA7. Ultrasound-based closed-loop control of implantable devices. Kostyantyn Shcherbina, Parag V. Chitnis (Dept. of BioEng., George Mason Univ., 4400 University Dr., 1G5, Fairfax, VA 22032, pchitnis@gmu.edu), and Samuel Sia (Biomedical Eng., Columbia Univ., New York, NY)

Implantable devices have a large potential to impact healthcare, but they are often made of biofouling materials that necessitate special coatings, rely on electrical connections for external communication, and require a continuous power source. We present an alternative platform, where an implanted biocompatible capsule can be wirelessly controlled by ultrasound to trigger the release of compounds. The capsule prototype was approximately 1 mm thick and consisted of a PDMS shell containing a co-polymer gel (NiPAAm-co-AAm). The co-polymer gel formulation was optimized so that it contracts when heated above body temperature (i.e., at 450C) to release compounds through an opening in the PDMS shell. This gel-containing capsule is free of toxic electronic or battery components. Ultrasound was employed for visualization (B-mode imaging), actuation (localized heating) and closed-loop control (ultrasound-based thermometry). This was achieved by integrating a diagnostic transducer coaxially within the central aperture of an spherically sectioned focused ultrasound (FUS) transducer. The combination of this gel-based capsule and compact ultrasound hardware can serve as a platform for triggering local release of compounds, including potentially in deep tissue, to achieve personalized and localized drug delivery. Illustrative results that show the ability to tailor release kinetics will be presented. [Work supported by NSF 1509748.]

3:00–3:15 Break

3:15

5pBA8. Development of an ultrasound tomography system: Preliminary results. Pedram Mojabi and Joe LoVetri (Elec. and Comput. Eng., Univ. of MB, 75 Chancellor's Circle, Winnipeg, MB R3T5V6, Canada, pedram.mojabi@gmail.com)

We are currently developing an ultrasound tomography system at the University of Manitoba. This system consists of eight circular rings of transducers so as to provide the potential to create two- and three-dimensional images. Thirty-two individual piezo-electric transducers are mounted in each ring, and each transducer can both transmit and receive ultrasonic waves. We utilize a frequency-domain inverse scattering framework to invert the measured ultrasound data so as to create quantitative images of the ultrasonic properties of the object being imaged. To this end, the measured data obtained from this system are calibrated before being inverted. For the calibration step, we discuss methods to accurately find the transducers' positions, properly convert the measured time-domain data into the frequency domain, and then minimize the discrepancy between the actual system and the simulated one through the use of calibration coefficients. Once this process is done, the calibrated frequency-domain data is processed by a frequency-domain inverse scattering algorithm. We utilize the Born iterative method (BIM) to invert this calibrated measured data to create two-dimensional images corresponding to ultrasonic properties of the object of interest [Mojabi & LoVetri, JASA, 2015]. Preliminary results of this BIM inversion are then presented and discussed.

3:30

5pBA9. Proposal of ultrasonic imaging of thermophysical property distribution *in vivo* by ultrasonic heating. Mai Morimoto, Yukako Tsujimoto, and Iwaki Akiyama (Medical Ultrasound Res. Ctr., Doshisha Univ., 1-3 Tatara-Miyakodani, Kyotanabe 610-0394, Japan, dmp1018@mail4.doshisha.ac.jp)

This study proposes a novel ultrasonic imaging of sound velocity change distribution by a short time ultrasonic heating. According to the bio-heat transfer equation, the sound velocity change due to ultrasonic heating depends on the volumetric heat capacity, attenuation coefficient, and the temperature dependence of the sound velocity. Thus, the resultant image is correspondent to the thermophysical property distribution. The sound velocity change is estimated by the time difference of echoes from the tissue between before and after heating. The feasibility of the proposed technique was studied by using a block of porcine fat and muscle exposed to ultrasound. A concave circular transducer of 5.2 MHz in resonance frequency was used for the measurement of sound velocity change. Ultrasound waves of 3.2 MHz in frequency was transmitted from a coaxial arranged concave ring transducer for heating. As a result, the sound velocity change of the porcine muscle is 1.3 m / s and the porcine fat is -2.3 m/s. Therefore, it is possible to identify the region of porcine muscle and fat, respectively. [This study was supported by MEXT-Supported Program for the Strategic Research Foundation at Priviate Universities, 2013-2017.]

3:45

5pBA10. Temperature measurement using backscattered ultrasonic power for non-invasive thermometry during HIFU ablation. David Melodelima and Victor Barrere (LabTAU - INSERM U1032, 151 cours Albert Thomas, Lyon 69003, France, David.Melodelima@inserm.fr)

The temperature dependence of ultrasonic tissue parameters has been reported extensively. In this work, the relationship between changes in ultrasound backscattered power and temperature during HIFU treatments was studied. An analytical model was developed based on attenuation, backscattered coefficient, velocity and thermal expansion to predict temperature changes. *In vitro* and *in vivo* tests were conducted in liver to confirm simulations. HIFU treatments were performed using a focused transducer working at 3 MHz. The radius of curvature and the diameter were both 70 mm. An ultrasound imaging probe working at 7.5 MHz was placed in the center of the HIFU transducer. Long exposure time (120 seconds) was used to observe smooth temperature increase from 37 to 70°C. Simulations predicted that backscattered power increased nearly logarithmically with temperature over the range from 37°C to 70°C. The model predicted a linear increase 10 dB. A linear increase 8 dB was measured in ultrasound backscattered power during experiments. The tissue temperature increase estimated using backscattered energy correlated well ($r = 0.79$) with temperature measurements performed using thermocouples. This linear relationship between changes in the backscattered energy and actual temperature was observed up to 70°C. Successful temperature estimation may allowed creating 2D temperature maps during HIFU treatments.

4:00

5pBA11. Color Doppler shear wave imaging and its application to breast imaging. Yoshiki Yamakoshi (Graduate School of Sci. and Technol., Gunma Univ., 1-5-1 Tenjin, Kiryu 376-8515, Japan, yamakoshi@gunma-u.ac.jp), Takahito Nakajima (Graduate School of Medicine, Gunma Univ., Maebashi, Japan), Mayuko Yamazaki, Ren Koda, and Naoki Sunaguchi (Graduate School of Sci. and Technol., Gunma Univ., Kiryu, Japan)

Imaging of continuous shear wave (CSW) induced by a small mechanical vibrator attached on tissue surface is low-cost and easy-to-use elastography. A problem of CSW is shear wave's reflection and refraction at tissue boundary, which degrade the accuracy of shear wave velocity estimation. We propose a novel shear wave elastography of CSW (Color Doppler Shear Wave Imaging: CD-SWI). Shear wave's propagation in soft tissue is reconstructed directly without adding any function to conventional color flow imaging instrument. Two conditions for shear wave frequency and displacement amplitude are required to obtain maps. However, they are not severe restrictions because frequency is chosen among several frequencies suited for elastography and the required displacement amplitude is a few tens of microns. Both Fourier analysis along time axis and a directional filter on wave-number vector space improve image quality. Accuracy and image quality of CD-SWI are evaluated by comparing with ARFI method. CD-SWI is applied to breast *in vivo* and shear wave propagation in breast ($f = 297$ Hz) is observed. Inherent shear wave propagation in soft tissue reconstructed by CD-SWI, such as diffraction and refraction around tissue with different stiffness as well as shear wave velocity, may give novel information in tissue characterization by shear wave.

5pBA12. Dynamic contrast specific ultrasound tomography. Libertario Demi, Ruud van Sloun, Hessel Wijkstra, and Massimo Mischi (Elec. Eng., Eindhoven Univ. of Technol., Den Dolech 2, Eindhoven 5612 AZ, Netherlands, l.demi@tue.nl)

Cumulative phase delay imaging (CPDI) is a modality recently introduced for contrast-specific ultrasound tomography. CPDI has already shown capable of imaging ultrasound contrast agents (UCAs) when working at pressures ($0.05 \leq MI \leq 0.2$) and frequencies (2.5-3 MHz) of interest for clinical applications. However, its ability of capturing UCA-kinetics was never shown. To this end, we imaged the passage of UCA-boluses through a dedicated gelatin flow-phantom. Each bolus resulted from a 5-mL injection with a 240-μL/L SonoVue® dilution. For imaging, the ULA-OP ultrasound research platform and an Esaote LA332 linear-array probe were employed. A frequency of 2.5 MHz and mechanical index (MI) equal to 0.07 were used. CPDI and Harmonic Imaging (HI) were simultaneously applied to each bolus-passage (in tomographic and echo-mode, respectively) to compare the two methods. Features commonly used to quantify UCA-kinetics were evaluated and used for the comparison: full-width-half-maximum (FWHM), wash-in-time (WIT), arrival-time (AR), and peak-time (PT). The obtained median absolute differences were equal to 0.625s, 0.25s, 0.1875s, and 0.06s for FWHM, AT, WIT, and PT, respectively. These results are encouraging, and open the way for the development of dynamic contrast-specific ultrasound tomography, possibly adding important features to multi-parametric ultrasound tomography of the breast, and improving breast cancer localization.

5pBA13. Null subtraction imaging: achieving super resolution by imaging with beam nulls. Jonathan Reeg and Michael L. Oelze (Dept. of Elec. and Comput. Eng., Univ. of Illinois at Urbana-Champaign, 405 N Mathews, Urbana, IL 61801, oelze@uiuc.edu)

Null subtraction imaging (NSI) uses multiple apodizations on receive to create images with low side lobes and super lateral resolution. The first apodization weight has a zero mean when summed across the aperture, placing a null in the beam at broadside. The second apodization simply adds a small DC offset to the first. The third apodization is the transpose of the second. An image created with the zero-mean apodization is then subtracted from the mean of images created with the second and third apodizations. Imaging tasks were performed using a 9L4 array connected to an Ultrasonix RP/SonixDAQ system on an ATS539 phantom containing wire, anechoic and hyperechoic targets. Images were constructed using NSI and compared with rectangular apodization. Image quality was assessed by estimating lateral resolution (-3-dB receive beamwidth), the mainlobe to sidelobe ratio (MSR) in dB and the contrast-to-noise ratio (CNR). NSI provided lateral resolution better than 25 times that of rectangular apodization. The MSR decreased by an average of 29 dB with NSI. However, because the speckle of NSI images become point like, the contrast of the contrast targets were reduced using NSI (e.g., CNR of hyperechoic targets was 1.35 and 0.36 for rectangular versus NSI).

5pBA14. Ultrasonic imaging of sum frequency components by crossed beam contrast echo method. Ayane Kihara, Kazumasa Kanai, and Iwaki Akiyama (Biomedical Information, Doshisha Univ., 1-3 Miyakodani, Tatara, Kyotanabe, Kyoto 610-0394, Japan, dmp1009@mail4.doshisha.ac.jp)

In a breast cancer diagnosis, the clinical usefulness of the contrast echo harmonic imaging has been reported. The authors proposed a contrast echo method by crossing the two ultrasonic beams to improve the accuracy. The microbubbles in the crossed region re-emit the secondary wave including the sum frequency components. It is capable of discriminating between the echo from microbubbles and the echo from tissue without microbubbles, since the sum frequency are not generated during the nonlinear propagation of primary wave outside the crossed region. Thus the contrast of the resultant image is improved. This study measured the sum frequency components in the echoes from the microbubbles of Sonazoid® in degassed water by using the ultrasonic imaging system with 64 channels. The ultrasonic wave of 5 MHz was transmitted from the array probe of central frequency of 6 MHz equipped with the imaging system. Since the beam formed by the array probe crossed the other beam of 2 MHz at the position 8 cm apart from the probe at 90 degrees, the value of sum frequency is the same as 7 MHz. 7 MHz components in the echo from the microbubbles in the crossed region was observed as 20 dB increased.

5pBA15. 3-D ultrafast ultrasound strain imaging to improve breast cancer detection. Gijs A. Hendriks, Chuan Chen, Hendrik H. Hansen, and Chris L. de Korte (Radiology and Nuclear Medicine, Radboud Univ. Medical Ctr., MUSIC 766, PO Box 9101, Nijmegen 6500HB, Netherlands, chris.dekorte@radboudumc.nl)

The automated breast volume scanner (ABVS) is used for breast cancer detection and consists of a linear transducer translating over the breast while collecting ultrasound data to reconstruct a breast volume. Although the ABVS shows high sensitivity, specificity remains a limitation resulting in high recall-rates. To improve ABVS specificity, we verified if it is feasible to implement 3-D strain imaging to discriminate between malignant (firmly-bound and stiff) and benign lesions (loosely bound and less stiff) based on lesion-to-surrounding-tissue connectivity (shear strain) and on stiffness (axial strain). Three phantoms containing loosely- (1) and firmly-bound (2) three times stiffer lesions were scanned twice by the ABVS, pre- and post-deformation (0.5 mm). Data was collected by plane-wave imaging to reduce acquisition times below one breath-hold. Displacements were calculated by cross-correlation from which axial and octahedral-shear strains were derived. Shear strains were distributed tightly and globally around the lesion for phantom (1) and (2), respectively. Using axial strain, the lesions could be clearly differentiated from the surrounding tissue in both phantoms (CNR_e 30; SNR_e 19dB). Therefore, we can conclude that it is feasible to implement 3-D strain imaging in the ABVS, and to improve its specificity by discriminating lesions by elastic properties.

Session 5pEA

Engineering Acoustics: Applied Acoustics II

Michael S. McBeth, Cochair
SSC Atlantic, SSC Atlantic/NASA Langley Research Center, 11 West Taylor Street, Mail Stop 207, Hampton, VA 23681

Hikaru Aono, Cochair
Mechanical Engineering, Tokyo University of Science, 6-3-1 Niijyuku, Katsushika-ku, Tokyo 125-8585, Japan

Contributed Papers

1:15

5pEA1. Acoustic field measurement of high-intensity ultrasound source system using acoustic waveguide for calibration of hydrophone. Shigeru Igarashi (Polytechnic Univ., 2-32-1 Ogawa-nishimachi, Kodaira-shi, Tokyo 187-0035, Japan, s-igarashi@uitec.ac.jp), Takeshi Morishita, and Shinichi Takeuchi (Toin Univ. of Yokohama, Yokohama, Kanagawa, Japan)

In recent years, high-intensity ultrasound equipments have been used. It is expected that nonlinear calibration and evaluation will be needed more often. It is necessary to develop a high-intensity ultrasound source that transmit a plane wave in far-field for calibration of hydrophone in order to evaluate high intensity acoustic field. We proposed an ultrasound source system using a cylindrical type acoustic waveguide and a concave type piezoelectric transducer. The focused ultrasound wave is emitted from the concave type transducer, propagated through the acoustic waveguide, and the output from that forms high-intensity acoustic beam. We simulated the output acoustic pressure of the proposed source system using axisymmetric 3D acoustic field simulation based finite element method, compared with that of a flat disk type transducer with the same aperture size as the acoustic waveguide, and the simulated data showed the acoustic pressure of our proposed system was higher than that of the flat disk type transducer with the same main beam width. We fabricated the experimental source system using the concave type piezoelectric transducer and the acoustic waveguide using closed cell sponge, and measured acoustic pressure distribution. The obtained results were nearly coincident with the simulated data

1:30

5pEA2. Effects of wing flexibility on sound characteristics of a four-wing flapping wing micro air vehicle. Hikaru Aono (Mech. Eng., Tokyo Univ. of Sci., 6-3-1 Niijyuku, Katsushika-ku, Tokyo 125-8585, Japan, aono@rs.tus.ac.jp), Yuta Ozawa (Graduate School of Eng., Tokyo Univ. of Sci., Tokyo, Japan), Makoto Yamamoto, Hitoshi Ishikawa (Mech. Eng., Tokyo Univ. of Sci., Tokyo, Japan), Chang-kwon Kang (Mech. and Aerosp. Eng., Univ. of Alabama, Huntsville, AL), Taku Nonomura (Inst. of Space and Astronautical Sci., Japan Aerosp. Exploration Agency, Sagamihara, Japan), and Hao Liu (Graduate School of Eng., Chiba Univ., Chiba, Japan)

The objective of this study was to understand the sound generation of flapping wings. In particular, the structural properties of the wings were varied to assess the effects of wing flexibility on the sound generation of flapping wings. A four-wing, hummingbird-inspired flapping wing micro air vehicle (FMAV) was considered. The half wing span was approximately six centimeters and the flapping frequency was approximately 25 Hz. The sound produced by the FMAV was measured using a 1/2 inches diameter microphone in an acoustic chamber under the quiescent flow condition. Time-averaged aerodynamic forces were measured by the force and torque sensor Nano17 (ATI Industrial Automation). Results revealed the characteristics of flapping-wing sound in terms of directivity, frequency response, and attenuation. Moreover, the degree of wing deformation led to a variation

in the frequency response corresponding to the maximum sound pressure level as well as time-averaged aerodynamic force, even when the flapping frequency was held constant in all the cases. All outcomes of this study would be helpful to control the sound generation of FMAVs and to develop new, silent FMAVs.

1:45

5pEA3. Battery outage and portable video recorders. Al Yonovitz (The Univ. of Montana, Dept. of Communicative Sci. and Disord., Missoula, MT 59812, al.yonovitz@umontana.edu), Joshua Yonovitz, Herbert Joe (Yonovitz & Joe, LLP, Dallas, Texas), and Brittany Galvin (The Univ. of Montana, Missoula, MT)

The advent of user-friendly compact video recorders with large storage capacities has made them ideal for recording investigations and interviews by law enforcement agencies. In addition, they are also useful for surreptitious recordings both by citizens and government services. When these recorders are powered by an internal battery, the battery recording may be interrupted when the battery discharges. The analysis of this premature stoppage of the recorded video file may be important in the determination of the authenticity of the recording for courtroom admissibility. The differentiation between a purposeful edit and the normal shutdown process was investigated with a series of such recordings using a variety of modern recorders and common smartphones. Recordings were made with the battery at 80 percent charge and allowed to deplete. The recorder was then immediately restarted as well as restarted while being immediately attached to permanent power. The resultant recordings were compared and analyzed for any extraneous information related to the shutdown process.

2:00

5pEA4. Detection of low intensity sound with laser Doppler anemometry. Michael S. McBeth (Space and Naval Warfare Systems Ctr. Atlantic, SSC Atlantic/NASA Langley Res. Ctr., 11 West Taylor St., M.S. 207, Hampton, VA 23681, m.s.mcbeth@ieee.org) and Robert Younts (Organic and Carbon Electronics Labs and Dept. of Phys., North Carolina State Univ., Raleigh, NC)

Laser Doppler anemometry (LDA) sound measurement systems have traditionally been designed to achieve point measurements of a sound field so they keep the laser path lengths through the water short to minimize the signal contribution from the acousto-optic effect and use relatively large beam separation angles to minimize the interference fringe volume. While the traditional LDA approaches provide reasonable sensitivity to the Doppler shifts from microparticles oscillating in a low frequency sound field, at high frequencies the sound measurement sensitivity depends on the optical path length difference (OPD). OPD increases due to the acousto-optic effect as the sound pressure is integrated over the entire path length. An LDA system using a small beam separation angle and long laser light paths through the water is shown to be capable of detecting low intensity sound pressure signals. Theoretical and experimental results are presented.

2:15

5pEA5. Optical displacement detection using multi-level diffraction gratings having no zeroth-order reflection. Randall P. Williams, Samuel K. Hord, and Neal A. Hall (Dept. of Elec. and Comput. Eng., The Univ. of Texas at Austin, 10100 Burnet Rd., Bldg. 160, Rm. 1.108, Austin, TX 78758, randy.williams@utexas.edu)

Displacement detection using optical interferometric techniques allows for low minimum detectable displacements which are unmatched by other displacement measurement methods. However, diffraction gratings commonly used in such applications can generate zeroth-order reflected beams which result in reduced sensor performance, packaging limitations, and laser instability. A new grating concept has been fabricated and tested which has the effect of eliminating the zeroth-order component by imparting a half-wavelength phase shift to a portion of the reflected light. The design criteria for zeroth-order beam elimination is illustrated using a simple model based on phasor arithmetic. Fabrication methods for producing these optical gratings, as well as comparisons between experimental test results and models, will be reported.

2:30

5pEA6. Sound source positioning using microphone array installed on a flying drone. Seongho Oh, Yeong-Ju Go, Jaehyung Lee, and Jong-Soo Choi (Aerosp. Eng., Chungnam National Univ., 99, Daehakro, Yuseong-gu, Daejeon 305-764, South Korea, osh09072@cnu.ac.kr)

This study presents an application of acoustic source localization using microphone array installed on a drone. Although recent technology of drone enables to record audio and video, there are still acoustic problems such as engine noise, propeller-induced flow noise, microphone directivity, etc. Such unwanted noise shows strong tonal characteristics with broad band noise contents. To acquire usable acoustic signal from array measurements while a drone is flying, we implemented acoustic signal conditioning and array processing. The approach to the problem is to adapt beamforming and TDOA techniques that provide localization capability of multiple sources and the position or direction of arrival of source, respectively. The array configuration and angle resolution are adjusted depending on the application requirements. We limit our attention to stable hovering condition to rule out position and heading error occurred by the drone motion. The experimental results of position estimation under a various noise conditions are discussed in the study.

2:45

5pEA7. Experimental evaluation of noise damping performance of in-duct perforated orifices with different geometric shapes. Dan Zhao (Aerosp. Eng. Div., Nanyang Technolog. Univ., Singapore, Singapore, Singapore), Chenzhen Ji, Nuomin Han (Aerosp. Eng. Div., Nanyang Technolog. Univ., School of Mech. and Aerosp. Eng., Singapore, Singapore, czji1000@gmail.com), and He Zhao (School of Energy and Power Eng., Jiangsu Univ. of Sci. and Technol., Zhenjiang, China)

In this work, 10 in-duct perforated plates are experimentally tested in a cold-flow pipe. These plates have the same porosities but different number N and geometric shaped orifices: 1) triangle, 2) square, 3) pentagon, 4) hexagon, and 5) star. The damping effect of these orifices is characterized by power absorption coefficient Δ and reflection one R from 100 to 1000 Hz. It is found that the orifice shape has little influence on Δ and R at lower frequency ($\omega_a/2\pi \leq 700$ Hz). However, as ω_a is increased, star-shaped orifice is shown to be with much lower Δ in comparison with that of other shaped orifices. For the in-duct plates with the same shaped orifices, increasing N does not lead to an increase of Δ at lower frequency ($\omega_a/2\pi \leq 900$ Hz). However, the orifice with the same shape and porosity but larger N is found to be associated with 20% more power absorption at approximately 970 Hz. Δ max is approximately 85% around 190 Hz, as Ma$\approx$0.29. The optimum Ma depends on the orifice shape. The present parametric measurements shed light on the roles of the number and geometric shapes of orifices and the flow parameters on its noise damping performance.

3:00

5pEA8. Omni-binaural recording system for 360° video. Takanori Nishino (Graduate School of Eng., Mie Univ., 1577 Kurimamachiya-cho, Tsu, Mie 5148507, Japan, nishino@pa.info.mie-u.ac.jp) and Kenta Niwa (Media Intelligence Labs., NTT, Musashino, Tokyo, Japan)

A binaural recording system for 360° video is developed and evaluated. The 360° images can be stitched by using multiple cameras. A binaural recording generally uses a dummy head microphone or binaural recording microphones; however, the orientation of the dummy head needs to be changed depending on the selected view. Binaural recording corresponding to multiple orientations of a dummy head is very difficult and sometimes impossible. Therefore, we propose a binaural recording system that can simultaneously record the sound for the entire corresponding 360° video. The proposed equipment is cylindrical in shape and has four hollows that correspond to the cavity of the concha on the side. Binaural signals corresponding to a desired direction are generated by a weighted average of the recorded signals. Three types of weights, i.e., linear, spline, and optimum weights, were examined. The degree of suitability between the video and generated binaural signals was evaluated using subjective tests. Nine subjects answered that all the weight conditions were acceptable and the proposed system was very interesting and entertaining.

3:15–3:30 Break

3:30

5pEA9. Acoustical design of the ELA4 launch pad for the Ariane 6 launcher. Hadrien Lambaré (CNES, CNES Direction des lanceurs 52 rue Jacques Hillairet, Paris 75612, France, hadrien.lambare@cnes.fr)

A new launch pad, ELA4, is going to be built in Kourou (French Guyana) for the future launcher Ariane 6. As the A64 version of the launcher will be more powerful at lift-off than Ariane 5, the exhaust ducts and the water injection systems have to be carefully designed, in order to be as effective as possible, and to reduce the acoustics on the payloads and the launcher. This has been done using the experience acquired and Ariane 5 and VEGA launch pads, as well as the understanding of the complex phenomena obtained by research activities. The general design have been validated and optimized by subscale test campaigns at the MARTEL test bench. Numerous configurations and parameters have been tested, including the geometry of the ducts and the launch table, the altitude of the launcher, the position and the mass flow rate and the spray type of the water injection systems. [These works have been done in collaboration between ESA, CNES, ASL, ONERA and PPRIME.]

3:45

5pEA10. Head mounted device for gesture recognition using acoustic time of flight sensor and 2D camera. Minchul Shin, Stephen Zeigler, Anton Gomez, and Christopher Reid (Mech. Eng., Georgia Southern Univ., Po Box 8045, Statesboro, GA 30465, mshin@georgiasouthern.edu)

The goal of this work is to design a 3D imaging system with acoustical transducers, 2D cameras and real-time signal processing to enable gesture sensing for a Head Mounted Device (HMD) design. The system offers enhanced user interfaces using vivid spatial information for use on mobile platforms such as wearable electronics. Through an intuitive user interface, hand gestures and finger shapes can provide easy acquisition of instruction as a sign language between devices and users. To achieve this intuitive gesture-based communication, key characteristics are required from imaging systems: light insensitivity, and high spatial density (high depth rate and x-y plane resolution). For gesture recognition, a variety of sensing methods using infrared, laser, and acoustics have been developed, but they have a critical issue: light disturbance. This device overcomes sunlight interference that limits 3D camera performance outdoors. Multiple microphones with resonant frequency 40 kHz are used as a depth sensor for 3D rendering or localization by using the TOF method. Eventually, motion tracking and localization via this 3D imaging system integrated with acoustic TOF sensor and RGB cameras allows communication using the dynamic and agile movements of users' hands and fingers and will serve as a major step toward expanding human-computer interactions.

4:00

5pEA11. Ultrasonic metal welding by longitudinal-torsional vibration source consisting of two transducers. Takuya Asami, Yusuke Higuchi, and Hikaru Miura (College of Sci. and Technol., Nihon Univ., 1-8-14, Kanda-surugadai, Chiyoda-ku, Tokyo 1018308, Japan, asami.takuya@nihon-u.ac.jp)

We study ultrasonic metal welding of dissimilar metals by using planar vibration locus. The planar vibration locus consists of the longitudinal vibration and the torsional vibration. In a previous study, we used an ultrasonic vibration source that contained a vibration converter comprising diagonal slits. The vibration source with diagonal slits was driven two frequencies in order to obtain a planar vibration locus. However, it was difficult to control individual vibration when using the vibration source with diagonal slits. In order to solve this problem, we have developed an ultrasonic longitudinal-torsional vibration source consisting of two transducers in which the longitudinal-torsional vibration can be controlled. The developed vibration source is consists of a longitudinal transducer and a torsional transducer attached to the both ends of a stepped horn with the length of two wavelength of the propagating vibration. This vibration source can generate planar vibration locus of substantially square. In this study, ultrasonic metal welding of dissimilar metals was studied by using the vibration source consisting of two transducers. As a result, high weld strength was obtained in the case of welding of the planar locus than welding of only longitudinal or torsional vibration loci.

FRIDAY AFTERNOON, 2 DECEMBER 2016

CORAL 4, 1:00 P.M. TO 4:30 P.M.

Session 5pID

Interdisciplinary: Topical Meeting on Data Science and Acoustics II

Matthew G. Blevins, Cochair
U.S. Army Engineer Research and Development Center, 2902 Newmark Drive, Champaign, IL 61822

Andrew Christian, Cochair
National Institute of Aerospace, 100 Exploration Way, Hampton, VA 23666

Hiroshi Sato, Cochair
Dept. of Information Technology and Human Factors, Natl. Inst. of Advanced Industrial Sci. and Tech., Tsukuba, Japan

Invited Papers

1:00

5pID1. Deep neural networks for learning classification features and generative models from synthetic aperture sonar big data. Johnny L. Chen (Appl. Res. in Acoust. LLC (ARiA), Culpeper, VA) and Jason E. Summers (Appl. Res. in Acoust. LLC (ARiA), 1222 4th St. SW, Washington, DC 20024-2302, jason.e.summers@ariacoustics.com)

Autonomous synthetic aperture sonar (SAS) imaging by unmanned underwater vehicles (UUVs) provides an abundance of high-resolution acoustic imagery useful for studying the seafloor and identifying targets of interest (e.g., unexploded ordnance or mines). Unaided manual processing is cumbersome as the amount of data gathered by UUVs can be enormous. Computer-vision and machine-learning techniques have helped to automate classification and object-recognition tasks, but often rely on hand-built features that fail to generalize. Deep-learning algorithms facilitated by emergence of graphics-processing unit (GPU) hardware and highly optimized neural-network implementations have recently enabled great improvements in computer vision. Autoencoders allow for deep unsupervised learning of features based on a reconstruction objective. Here, we present unsupervised feature learning applied to seafloor classification of SAS images. Deep architectures are also capable generative models. We illustrate this with generative networks that are capable of generating realistic SAS images of different seafloor bottom types. Deep models allow us to construct algorithms that learn hierarchical and higher-order SAS features, which promise to improve automatic target recognition (ATR) and aid operators in processing the large data volumes generated by UUV based SAS imaging. [Work supported by the Office of Naval Research.]

1:20

5pID2. Acoustic and seismic time series analysis using ensemble empirical mode decomposition. Gopu R. Potty and James H. Miller (Dept. of Ocean Eng., Univ. of Rhode Island, 115 Middleton Bldg., Narragansett, RI 02882, potty@egr.uri.edu)

Empirical mode decomposition (EMD) is an adaptive or data-driven time series analysis technique ideally suited to investigate non-stationary signals. EMD decomposes the signal into fast and slow oscillations called intrinsic mode functions (IMFs). Ensemble EMD (EEMD) methods have been developed to alleviate the mode-mixing phenomenon present in the EMD technique. Various real signals will be analyzed using these methods to investigate whether the IMFs correlate to acoustic modes. These real signals include underwater acoustic signals from broadband sources, Scholte and Rayleigh wave signals and music signals. These underwater acoustic signals and the seismic signals exhibit multi-modal structure. In addition these modes are also dispersive in nature. Accurate resolution of these modes in the time-frequency plane is critical for the estimation of medium properties via inverse schemes. Finally, the EEMD technique will be used to

3423   J. Acoust. Soc. Am., Vol. 140, No. 4, Pt. 2, October 2016

5th Joint Meeting ASA/ASJ   3423

investigate the complex dynamic structure of the pitching structure in South Indian Classical music. In all these cases investigated the mapping of the signal into the time-frequency plane reveal distinct features representing the modal dispersion and pitching structure, respectively. EEMD will be used to extract these features accurately with high resolution. [Work supported by Office of Naval Research.]

1:40

5pID3. Deep learning for unsupervised feature extraction in audio signals: Monaural source separation. Edward T. Nykaza (ERDC, 2902 Newmark Dr., Champaign, IL 61822, edward.t.nykaza@erdc.dren.mil), Arnold P. Boedihardjo (ERDC, Alexandria, Virginia), Zhiguang Wang, Tim Oates (Comput. Sci. and Elec. Eng., UMBC, Baltimore, MD), Anton Netchaev, Steven L. Bunkley (ERDC, Vicksburg, MS), and Matthew G. Blevins (ERDC, Champaign, IL)

Deep learning is becoming ubiquitous; it is the underlying and driving force behind many heavily embedded technologies in society (e.g., search engines, fraud detection warning systems, and social-media facial recognition algorithms). Over the past few years there has been a steady increase in the number of audio related applications of deep learning. Recently, Nykaza *et al.* presented a pedagogical approach to understanding how the hidden layers recreate, separate, and classify environmental noise signals. That work presented some feature extraction examples using simple pure tone, chord, and environmental noise datasets. In this paper, we build upon this recent analysis and expand the datasets to include more realistic representations of those datasets with the inclusion of noise and overlapping signals. Additionally, we consider other related architectures (e.g., variant-autoencoders, recurrent neural networks, and fixing hidden nodes/layers), explore their advantages/drawbacks, and provide insights on each technique.

2:00

5pID4. Comparison between multilevel and traditional logistic regression analysis of laboratory indoor annoyance caused by shaped sonic booms. Jonathan Rathsam (NASA Langley Res. Ctr., Hampton, VA), Noah H. Silbert (Commun. Sci. and Disord., Univ. of Cincinnati, French East 344, 3202 Eden Ave., Cincinnati, OH 45267, silbernh@ucmail.uc.edu), Kathryn Ballard (NASA Langley Res. Ctr., Hampton, VA), and Jasme Lee (Dept. of Statistics and Biostatistics, Rutgers Univ., Piscataway, NJ)

In the 1970s commercial supersonic flight was forbidden over communities because conventional N-wave sonic booms caused excessive annoyance. Through recent advances in aircraft design techniques, conventional sonic booms are being modified into quieter, shaped sonic booms. As noise regulators consider allowing shaped booms over communities, NASA is providing data and expertise. Recently, a laboratory study quantified the additional annoyance caused by directly felt structural vibrations (J. Rathsam and J. Klos, "Vibration penalty estimates for indoor annoyance caused by sonic boom," J. Acoust. Soc. Am. **139** (4), 2007 (2016)). To confirm experimental results and complement the traditional logistic regression analysis, a multilevel logistic regression model was fit to individual-level data. The multilevel analysis allows us to simultaneously model individual- and group-level factors. As a result, it is straightforward to test for relationships between annoyance and individual-level factors such as noise sensitivity.

2:20–2:40 Break

2:40

5pID5. Anomaly detection and other practical considerations for estimating acoustical metrics from time series data. Daniel J. Mennitt (Elec. and Comput. Eng., Colorado State Univ., 1201 Oakridge Dr., Ste. 100, Fort Collins, CO 80525, daniel_mennitt@partner.nps.gov) and Kurt Fristrup (Natural Sounds and Night Skies Div., National Park Service, Fort Collins, CO)

The scope of environmental acoustic monitoring and rate of data collection are growing rapidly. These increases in the quantity of information have elevated the necessity of detecting anomalous data and the difficulty of doing so. Analysis of contaminated data leads to incorrect results, including biased parameter estimation and flawed model selection. Censoring of data requires strong justification, as the loss of information can lead to these same problems. For acoustics, this issue is compounded by the widespread use of time-average sound levels (L_{eq}), which are especially sensitive to anomalous measurements. This talk will discuss applications of anomaly detection to acoustic pressure time series data for the calculation of long-term metrics. There is no single universally applicable or generic approach to data cleaning and preparation, which often consumes a disproportionate amount of time and effort relative to interpretation of results. Examples drawn from terrestrial and aquatic acoustical monitoring data sets of varying resolution and associated metadata will provide context for issues that compromise data quality, anomaly detection methods, and mitigation efforts.

3:00

5pID6. The effect of sample rate reduction and audio compression on noise metric accuracy and statistical learning classifier performance. Matthew G. Blevins, Edward T. Nykaza (U.S. Army Engineer Res. and Development Ctr., 2902 Newmark Dr., Champaign, IL 61822, matthew.g.blevins@usace.army.mil), and Anton Netchaev (U.S. Army Engineer Res. and Development Ctr., Vicksburg, MS)

With the increasing availability of low-cost and portable noise monitors, the amount of captured data can reasonably be predicted to grow exponentially. The financial and computational expense of portable data storage, processing hardware, and wireless or cellular upload bandwidth, however, necessitate data compression techniques to ensure feasible continuous operation. Sample rate reduction causes a loss of high frequency components while compressed formats (such as mp3) remove parts of the signal that are inaudible to humans, but may be important to statistical learning algorithms. The degree to which these data reduction techniques degrade the accuracy of metrics calculated from the signals (e.g. peak level, equivalent level, sound exposure level, etc.) and the performance of statistical learning classifiers that use the metrics and/or band levels is examined. Over 20,000 recordings from noise monitors located on military installations are systematically down-sampled and compressed to use in the analysis. The root-mean-square error for each of over 40 noise metrics and the mean decrease in accuracy for a statistical learning classifier are presented.

Contributed Papers

3:20

5pID7. Facilitation of bird call classification using self-organising maps—A pilot study on bird calls in New Zealand. Tsutomu Miyoshi (Dept. of Media Informatics, Faculty of Sci. and Eng., Ryukoku Univ., Yokotani 1-5, Seta Oe-cho, Otsu-shi, Shiga 520-2194, Japan, mijosxi@rins.ryukoku.ac.jp) and Yusuke Hioka (Mech. Eng., Univ. of Auckland, Auckland, New Zealand)

Automatic detection and classification of bird calls from audio recordings attract interests of various users from amateur bird watching hobbyists to professional biologists and park rangers studying biodiversity and conservation. Previous studies utilize machine learning techniques many of which require supervised learning; therefore, they cannot detect bird calls that are not included in the supervisory data. Their performance also severely degrade when a segment of a bird call is overlapped by other birds' calls or contaminated with various ambient noise. This study proposes an alternative approach for bird call classification, which is facilitated by an algorithm based on the self-organising maps (SOM). Since the SOM is an unsupervised learning algorithm, which classifies similar input features into neighboring clusters, it is expected that segments of different bird species will be classified into clusters located apart while segments with multiple bird calls and/or noise will lie down on the clusters in the middle. The classified segments are then labelled by measuring the distance between the feature vectors in each cluster and that of known bird call segments given a priori. The performance of the proposed method is evaluated by applying the algorithm to data recorded in a national park in New Zealand.

3:35

5pID8. Noise injection for acoustics fault sample expansion. linke zhang (School of Energy and Power Eng., Wuhan Univ. of Technol., 1178 Heping Ave., Wuhan, Hubei 430063, China, lincol_zhang@126.com) and NA Wei (Inst. of Acoust., Nanjing Univ., Nanjing, China)

Sample expansion is an effective approach to resolve the incomplete samples problem in acoustics fault source identification. Noise injection is applied to acoustics fault sample expansion by mixing noise signals with different signal-to-noise ratio in real samples to obtain expansion samples, which enhances training data. Experimental results demonstrate the effectiveness of this proposed approach.

3:50–4:30 Panel Discussion

FRIDAY AFTERNOON, 2 DECEMBER 2016 CORAL 3, 1:00 P.M. TO 4:00 P.M.

Session 5pMU

Musical Acoustics: General Topics in Musical Acoustics II (Poster Session)

Andrew C. Morrison, Chair
Natural Science Department, Joliet Junior College, 1215 Houbolt Rd., Joliet, IL 60431

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

Contributed Papers

5pMU1. Real-time gaze-tracking performance under synchronous and asynchronous auditory feedback conditions. Andrew Lucila (Psych., William Paterson Univ., 780 Allwood Rd., Clifton, NJ 07012, Drewlucila@gmail.com), Franklin Roque (Psych., William Paterson Univ., Paramus, NJ), Michael Morgan, and Michael S. Gordon (Psych., William Paterson Univ., Wayne, NJ)

Visual attention and manual coordination were tested in pianists and other instrumentalists tasked with reading baroque music. Musicians were assessed in their ability to perform a single melodic line relative to a two-clef, three voice score with bimanual coordination (an extension from past research). In this study, musicians were asked to read and perform music under three auditory feedback conditions: synchronous, delayed (one-beat behind), and prelayed (one-beat ahead). Real-time gaze-tracking analysis was conducted, in addition to measures of performance, under each auditory condition. Results from these analyses indicated several trends. Gaze-tracking measures were taken from visual fixations and regressions during performance. These measures indicated that musicians had more difficulty in the asynchronous auditory feedback conditions relative to the synchronous condition. Particularly the prelay auditory feedback condition tended to cause an increase in the number of visual fixations per measure and required more effortful gaze. This effect was found to be consistent across different instrumental performances. In addition, indications of cognitive effort increased in the bimanual conditions. These results indicate the challenge musicians face when reading music in conditions of higher cognitive and manual load.

5pMU2. A study on acoustic feature representing breathiness of singing voice based on vocal-fold vibration modeling. Masahiro Itou, Hideki Banno, and Kensaku Asahi (Graduate School of Sci. and Technol., Meijo Univ., 1-501 Shiogamaguchi, Tempaku-ku, Nagoya-shi, Aichi-ken 468-0073, Japan, 153430003@ccalumni.meijo-u.ac.jp)

There are some techniques expressing emotion in singing. One of the techniques used by some professional singers is controlling breathiness of uttered voice intentionally. However, this breathiness control can make it difficult for an amateurish singer to keep pitch and pressure of uttered voice steady. In this publication, we verify what kind of feature is appropriate for representing breathiness of singing voice by converting the feature of normal voice into that of breathy voice. Extracted features for conversion are aperiodicity index and STRAIGHT spectrum ("voice quality parameters" below) obtained through STRAIGHT analysis of recorded vowel singing whose pitch, pressure and breathiness are constant and under several conditions. We have tested two conversion methods; one compensates voice quality parameters of normal voice statically so that the averaged parameters of generated voice are equivalent to those of high-breathiness voice, another exchanges voice quality parameters of normal voice totally with those of high-breathiness voice. It was found that the generated sound by the latter method is perceived to have higher quality than that by the former method. This result implies that features to represent breathiness properly must include not only static information but also dynamic information of breathy voice.

5pMU3. An application of multiple-stage non-negative matrix factorization to music analysis. Kenko Ota, Takuya Ichigo, and Takahiro Shirasawa (Nippon Inst. of Technol., 4-1 Gakuendai, Miyashiro, Minami-saitama 345-8501, Japan, otakenko@nit.ac.jp)

Non-negative matrix factorization (NMF) is an analysis technique for matrix with non-negative elements. NMF decomposes a sound spectrogram X obtained by a short-time Fourier transform into the product of two nonnegative matrices H and U. The matrix H represents the bases of X and the matrix U represents the activation gains of H. As an extension of NMF, multiple-stage NMF has been proposed in order to represent the hierarchy of data. This research focuses on the hierarchy of music data, and attempts to apply multiple-stage NMF to music analysis. For simplicity, sounds treated in this research are assumed as follows: single tone consists of fundamental tone and harmonic overtones and a chord consists of some single tones. According to this assumption, observed sound spectrogram can be decomposed into three stages. In the first stage, the basis matrix is approximated by a Gaussian distribution in order to represent a spectrum of fundamental tone and harmonic overtones. In the second stage, the basis matrix consists of the set of single tone with fundamental tone and harmonic overtones. And in the third stage, the basis matrix consists of the set of chord with some single tones. Evaluation was carried out by MIDI data.

5pMU4. An acoustic analysis of *Arirang*, a representative folk song of Korea. Sang-Bum Park (Dept. of Information and Technol. Eng., Soongsil Univ., Seoul, South Korea), Myung-Sook Kim (Dept. of English, Soongsil Univ., Sando-ro 369, Dongjak-gu, Seoul, Korea, Seoul 06978, South Korea, kimm@ssu.ac.kr), and Myung-Jin Bae (Dept. of Information and Technol. Eng., Soongsil Univ., Seoul, South Korea)

Arirang, a representative folk song of Korea, became inscribed on the Representative List of the Intangible Cultural Heritage of Humanity program by UNESCO in 2012. Since then, the Korean government as well as some Korean private sector companies have been promoting the song and its unique feelings in melody and lyrics drawn from Korea's own historical backgrounds. This study presents an acoustic analysis of *Arirang* from a psychoacoustic point of view, first concerning its rhythm/tempo and second concerning its acoustic characteristics found in certain vowels and consonants in lyrics. Since *Arirang* has so many versions according to different regional origins in Korea, *Arirang* can produce contradictory feelings to its listeners depending on its version as well as its tempo. In general, when played slowly, the listeners may feel sad and compassionate, but when played fast the listeners may feel cheerful and even excited. As far as its lyrics are concerned, certain vowels and consonants evoke special feelings: first, the vowel /a/, which appears so many times in the song as in its title, has higher fundamental frequency in *Milyang Arirang* (200 Hz) than in *Jindo Arirang* (400 Hz), evoking a variety of emotions; second, the consonant /r/, as in its title, covers a wide range in amplitude with many variable frequency bandwidths than other consonants and thus produces a cheerful feeing.

5pMU5. Recognizing musical melodies: Effects of audio and visual representations of emotion. Alejandro L. Ataucusi (Psych., William Paterson Univ., 909 Florida Grove Rd., Keasbey, NJ 08832, ataucusia@student. wpunj.edu) and Michael S. Gordon (Psych., William Paterson Univ., Wayne, NJ)

Memory for music was tested with a set of emotionally valenced melodies encoded with and without an accompanying visual scene. The music consisted of original melodies that were normed for their emotional valence. During the training phase, participants were presented a set of happy, neutral, and sad musical melodies. Half of the melodies were presented with an affective, visual stimulus, depicting a scene consistent with emotional expression of the melody. After a distraction period, participants were tested in an audio-only recognition task. The recognition task included a mix of melodies from the training phase (with equal representation of audiovisual and audio conditions and each of the affective conditions) and emotionally-matched novel solo guitar melodies. Using signal detection analyses participants were found to be more sensitive to stimuli for which they had been trained in the Auditory-only conditions than in the Audiovisual conditions. Sensitivity also tended to be highest with the neutral affective melodies and

lowest with the sad melodies. These findings show the potential for distraction with visual imagery to musical memory formation—despite the consistency of that emotional content to the melody. Moreover, less emotionally valenced music tended to be more effectively encoded.

5pMU6. Acoustic analysis on vocal voice of Lee Nan-Young, a legendary singer of Korea. Myungsook Kim (English, Soongsil Univ., Sangdo-ro 369, Seoul 06978, South Korea, kimm@ssu.ac.kr) and Myung-Jin Bae (Information and Technol. Eng., Soongsil Univ., Seoul, South Korea)

Lee Nan-Young, a legendary female popular singer of Korea, made her debut in 1934 and enjoyed a tremendous popularity among Korean people in the 1930-40s. She was a special comfort to Korean people who were suffering from colonial exploitations by the Japanese government during that time. To celebrate her 100th birthday, we analyzed her vocal recordings in order to find its unique characteristics that set her apart. Findings can be summarized as follows: first, her voice has harmonics found around 4500 Hz and even higher, which is three times higher than average; second, she changes tones from 250 Hz to 1500 Hz (2.5 octaves), producing an attractive plaintive voice; third, she gives many short cut-offs in low tones but produces rich resonant voice in high tones, practicing a variety of singing methods; fourth, she has a command of producing all-frequency-ranged voice including low, middle, and high frequency and makes rich vibrations resulting in deep and touching voice; last, her voice naturally and smoothly links with a variety of musical instruments accompanying her songs. Based on these findings, we conclude that she deserves to be recognized as one of the best popular singers who have ever performed.

5pMU7. Study on specific musical instrument sound extraction by clustering of nonnegative matrix factorization bases with linear predictive coding mel-cepstrum. Saki Umeda, Hideki Banno, and Kensaku Asahi (Meijo Univ., TenpakukuShiogamaguchi1-501, Nagoya 468-8502, Japan, 163430005@ccalumni.meijo-u.ac.jp)

Nonnegative matrix factorization (NMF) is one of a technique for audio source separation. It separates input audio signal into a set of basis spectra and a set of intensity trajectories of the basis spectra. We utilize this technique to extract particular musical instrument sound from musical audio signal. However, the resultant separation by NMF is known to be unrelated to instrumentation of the audio signal. Therefore, clustering technique by using a spectral distance measure is applied to the basis spectra so that the clustered signal corresponds to the audio signal of particular musical instrument. In this study, we investigate the optimal number of bases and the optimal distance measure for clustering. Experiments using audio signals that include two musical instrument sounds and employing mel-cepstrum, linear predictive coding (LPC) cepstrum and LPC mel-cepstrum as a distance measure, showed that LPC mel-cepstrum tends to classify more properly than the others do. This tendency is caused by the fact that the resultant spectrum obtained by LPC mel-cepstrum keeps some important peaks to classify musical instrument sounds properly. Future work includes integrating LPC mel-cepstrum into a distance measure used in NMF algorithm inside to fit with extraction of particular musical instrument sound.

5pMU8. A simple model of cross-fingering explaining both pitch flattening and sharpening. Seiji Adachi (Dept. of Acoust., Fraunhofer Inst. for Bldg. Phys., Nobelstr. 12B, Stuttgart 70569, Germany, seiji.adachi@ibp. fraunhofer.de)

Cross-fingering is a technique of playing woodwind instruments in which one or more tone holes are closed below the first open hole. It usually yields a pitch lower than that played with normal fingering. There is, however, an exception resulting in pitch sharpening. This paper proposes understanding these pitch bending phenomena in a unified manner with a model of two coupled mechanical oscillators, each of which represents a bore above or below the open hole. This coupled system has two resonance frequencies $\omega_{\pm}$, which are respectively higher and lower than those of the upper and lower bores ω_U and ω_L. The $\omega_{\pm}$ differ even if $\omega_U = \omega_L$. The normal effect of cross-fingering, i.e., pitch flattening, corresponds to excitation of the ω_--mode, which occurs when $\omega_L \geq \omega_U$ and the admittance peak of the ω_--mode is higher than or as high as that of the ω_+-mode. Excitation of

the ω_+-mode yields pitch sharpening. This occurs when $\omega_L \leq \omega_U$ and the peak of the ω_+-mode becomes sufficiently high. With an extended model having three degrees of freedom, pitch bending of the recorder played with cross-fingering in the second register can also be successfully explained.

5pMU9. Relation between violin timbre and harmony overtone. Masao Yokoyama, Yoshiki Awahara (Information Sci., Meisei Univ., 2-1-1 Hodokubo, Hino 191-8506, Japan, masao.yokoyama@meisei-u.ac.jp), and Genki Yagawa (Tokyo Univ., Tokyo, Tokyo, Japan)

The timbre of violins has been studied by several researchers from various points of view including the structure, the acoustic characteristic, the chemical composition of the varnish and the acoustic radiation. Although many of them have mentioned that the Stradivari's violin gives the most beautiful timbre, the reasons of which have not been clarified yet. We have studied the timbre of about 30 violins from old ones to new ones, namely, the relation between the harmonic overtones and the expression words, which the audience receives from the sound of violin. But, we have not been successful in clarifying how the structure of overtone is related with the feeling of the listeners of sound such as "rich", "blight" and "soft" as the timbre of violin. In this paper, we have analyzed the change of overtone structure with the difference of violinist's performance. For instance, the power (dB) of sound at the frequency area over 2000 Hz played in fortissimo or near a bridge like sul ponticello was approximately 10% to 20% higher than that played in pianissimo or on finger board such as sul tast. We also found that the power of non-harmonic overtone of the sound which was played in the "rich" or "soft" expression was high at the frequency area over 2000 Hz and the kurtosis of the peak of harmonic overtone was small.

5pMU10. Effect of hardening piano hammer felt on piano sound. Shingo Takaku (Hamamatsu Gakuin Univ. / Shizuoka Univ., 3-2-3 Nunohashi,-Naka-ku, Hamamatsu 432-8012, Japan, takaku@hgu.ac.jp) and Yosuke Tatekura (Shizuoka Univ., Shizuoka, Japan)

To clarify the effect of hardening the hammer felt of a piano on its sound, experiments evaluating the sound quality of the piano were performed using a new piano hammer. The adjustment of a piano hammer is necessary to obtain the desired tone. The elasticity and hardness of the hammer are also recognized as indicators of the condition of the instrument. To maintain the condition of piano hammers, piano craftsmen make adjustments using various tools and chemicals while listening to the resulting change in the sound after each alteration. Hammer adjustment is often performed based on intuition and experience. In this study, the quantity of the hardening agent was gradually increased, and a sound quality experiment was performed using a new Renner hammer. A low frequency of 58.27 Hz (A#1) was selected as the piano note investigated in this study because it is more affected than other frequencies by the hardening of the hammer. The results of this study indicate that the optimal amount of hardening agent needed to achieve a desired tone and desired tone should be determined before the agent is applied to the hammer felt.

5pMU11. Warbling in steadily bowed viol strings. D. Murray Campbell (School of Phys. and Astronomy, Univ. of Edinburgh, James Clerk Maxwell Bldg., Mayfield Rd., Edinburgh EH9 3JZ, United Kingdom, d.m.campbell@ed.ac.uk), Patsy Campbell (Edinburgh College of Art, Univ. of Edinburgh, Edinburgh, United Kingdom), and Jim Woodhouse (Eng. Dept., Cambridge Univ., Cambridge, United Kingdom)

The timbre of a note played on a bowed string instrument depends on a number of performance parameters, including bow pressure, bow speed, and distance of the bowing point from the bridge. When these parameters are held constant during the bowing of a note played without vibrato the radiated sound usually has a stable timbre, with a frequency spectrum which is unchanging apart from small random fluctuations. For certain combinations of instrument and bow this timbral stability is disrupted, leading to a periodic modulation of the timbre which can be described as a "warble." The warbling effect was first noticed on a large bass viol, but it can also be found on other bowed string instruments including the cello. Unlike the well known "wolf" phenomenon, the warbling effect is not associated with a particular played pitch on the instrument, but occurs over a wide range of pitches. Experimental studies on bass viols have shown that the warble is a modulation of upper harmonics of the played note, with a modulation frequency of a few hertz. Possible mechanisms which could explain the generation of this modulation are discussed.

5pMU12. The effect of attack time and release time of the sound source on the sense of listener envelopment. Toru Kamekawa and Atsushi Marui (Musical Creativity and the Environment, Tokyo Univ. of the Arts, 1-25-1, Senju, Adachi-ku, Tokyo 120-0034, Japan, kamekawa@ms.geidai.ac.jp)

The effect of attack time and release time of the sound source on the sense of Listener Envelopment (LEV) was investigated. From the results of a previous experiment, it was found that there were at least two kinds of sense in LEV, namely, spatial continuity between front and back, and expanse of lateral direction. In this experiment, authors conducted pairwise comparison tests for each sense using simple sine wave stimuli changing four kind of attack time and three kind of release time. Furthermore these short sine waves were convolved with four channel room impulse responses and played back form five channel loud speakers. The result of pairwise comparison showed that the sense of spatial continuity between front and back was related to first attack time and another sense of expanse of lateral direction was related to slower attack time.

5pMU13. Acoustic and perceptual differences between novice and professional music theatre singers. Lynn M. Maxfield (National Ctr. for Voice and Speech, Univ. of Utah, 136 S Main St., Ste #320, Salt Lake City, UT 84101, lynn.maxfield@utah.edu) and Brian Manternach (Dept. of Theatre, Univ. of Utah, Salt Lake City, UT)

Research examining contemporary commercial music (CCM) styles of singing has increased significantly over the last ten years. While acoustic analysis has helped define which characteristics define various vocal genres, discrepancy still exists in how those acoustic characteristics are perceived, described, and evaluated. The current study recorded novice and professional musical theatre singers performing belt, legit, and mix vocal samples. Three acoustic analyses were applied to the excerpted recordings from each singer using the Praat voice and speech analysis software: The spectral slope of the long term average spectrum (LTAS) was calculated, as well as the noise to harmonic ratio (NHR), and the dominant harmonic (1st, 2nd, or 3rd) was noted for each sample. Results were compared across pitch, style (belt v. legit), and training level (professional v. novice). Finally, raters listened to the recordings and rated each sample on the basis of style (belt v. legit), roughness (rough v smooth), and tone quality (brassy v flutey, and bright v dark). A 120mm liechert scale was provided to each rater for each of the rating variables. Results of the perceptual assessments were compared with acoustic measures to elucidate discrepancies in how ccm characteristics are perceived.

5pMU14. Vowel duration dependent hidden Markov model for automatic lyrics recognition. Shohei Awata, Shinji Sako, and Tadashi Kitamura (Comput. Sci., Nagoya Inst. of Technol., Gokisocho,Showa-ku Nagoya-shi,Aichi, Nagoya 466-8555, Japan, awata@mmsp.nitech.ac.jp)

Recently, due to the spread of music distribution service, a large amount of music is available on the Internet. Accordingly, it is generally increasing the demand of music information retrieval (MIR). In the field of MIR research, there are several researches to extract meaningful information from music audio signals. However, automatic lyrics recognition is still a challenging problem because the variation of singing voice is much larger than that of speaking voice and a large database of singing voice is not available. In the relevant study, lyrics recognition was performed by extending the framework of speech recognition using hidden Markov model (HMM). However, accuracy rate was not sufficient. To recognize singing voice precisely, one promising approach is utilizing musical features. This study considers the task of recognizing syllable from a cappella singing voice. To respond to the variation of the length of a phoneme, we construct the duration dependent HMM. A large database of singing voice is essential for training the acoustic model. We use synthetic singing voice by HMM based singing voice synthesis system to solve the lack of the database of a cappella singing voice. We confirmed the effectiveness of our method.

3427   J. Acoust. Soc. Am., Vol. 140, No. 4, Pt. 2, October 2016

5th Joint Meeting ASA/ASJ   3427

5pMU15. Sound source separation and synthesis for audio enhancement based on spectral amplitudes of two-channel stereo signals. Masayuki Nishiguchi, Ayumu Morikawa, Kanji Watanabe, Koji Abe, and Shoichi Takane (Dept. of Electronics and Information Systems, Akita Prefectural Univ., 84-4 Ebinokuchi, Tsuchiya, Yurihonjo, Akita 015 0055, Japan, nishiguchi@akita-pu.ac.jp)

A sound source separation algorithm based on the spectral amplitudes of 2-channel signals has been developed for the up-mixing playback of 2-channel stereo. Short-term Fourier transforms (STFT) of the signals on the left and right channels are first calculated. The coefficients of the discrete Fourier transform (DFT) are used to calculate the ratio of the spectral amplitudes of the left and right channels, which is termed the channel level difference (CLD). The DFT coefficients are then divided into multiple groups on the basis of the CLD, with each group representing a separated sound source. The signal-to-distortion ratio (SDR)is used to evaluate the signal separation performance. It was found that a rough estimate of the CLD threshold yielding the best SDR could be obtained by cross-correlating the separated sounds. For playback on a headset, each separated signal is convoluted with head-related transfer functions (HRTF) that represent the direction of that particular sound source. Subjective listening tests showed that the sound synthesized by this method is more realistic than that synthesized with HRTFs that represent only left and right speakers.

5pMU16. Study on relationship between subjective reproducibility of individuality and distance measure for vibrato of singing voice. Chifumi Suzuki, Hideki Banno, Kensaku Asahi (Graduate School of Sci. and Technol., Graduate School of Meijo Univ., 1-301 Shiogamaguchi Tenpaku-ku, Nagoya 468-8502, Japan, 123441501@ccalumni.meijo-u.ac.jp), and Masanori Morise (Univ. of Yamanashi, Koufu, Japan)

In this research, a distance measure between two vibratos had been developed to evaluate singing style. This paper verifies whether the measure properly represents subjective distance by the following two subjective experiments. In these experiments, synthetic sounds whose vibratory amplitude computed from a target singer's singing is variable were used, and listeners were asked to answer reproducibility of individuality of the target singer for the synthetic sound. In the first experiment, listeners were asked to remember vibrato of the target singer by listening to the singing. As a result, it was found that not only synthetic sounds showing small distance to the target singer in the distance measure but also those not showing small distance produced higher subjective reproducibility, as long as the sound whose vibratory amplitude is large, i.e., includes *deformation* in the F0 trajectory. To confirm whether this tendency is caused by the experimental procedure, the second experiment that listeners were compared with the reference synthetic sound was conducted. As a result, it was found that the results were mostly the same as those of the first experiment, and the vibrato rate affects the subjective reproducibility more strongly, as the magnitude of vibratory amplitude becomes large.

5pMU17. Tempo estimation for acoustic signal of tempo variation music. Sota Okada (Dept. of Media Informatics, Graduate School of Ryukoku Univ., 1-5, Yokoe, Oe-Cho, Otsu, Shiga 520-2153, Japan, xkre3x@gmail.com) and Masanobu Miura (Dept. of Media Informatics, Ryukoku Univ., Otsu, Shiga, Japan)

This study proposes a tempo estimation method for acoustic signal of tempo variation music. Previous studies observe the power spectrum of envelope curve for its acoustic signal, where its length must be sixty seconds or longer for the purpose of keeping the tempo resolution as 1 bpm. The method then picked the highest power up as estimated tempo. The music with tempo variation that is occurred within the sixty seconds (for example ten seconds) were, however, hard to estimate. In order to estimate tempo for each time of such tempo variation music, the conventional zero-padding method is carried out when performing FFT. Tempo value is therefore estimated under high accuracy yet the estimation errors have not discussed so far. Accordingly, the optimum ratio of zero-padding was investigated. As a result, the optimized ratio of acoustic signal and 0 in the zero-padding

is.125—.875 %. Furthermore, tempo value on each one second is estimated for tempo variation music that has orchestral rubato expression. The estimated accuracy of presented method was evaluated by comparing estimated with hand-labeled tempo.

5pMU18. Drum sound onset detection based on class separation using deep neural network. Takahiro Kuriwaki, Takanori Nishino, and Hiroshi Naruse (Graduate School of information Eng., Mie Univ., 1577 kurimamachiya-cho, Tsu, Mie 514-8507, Japan, kuriwaki@pa.info.mie-u.ac.jp)

This study reports an onset detection method for the snare drum sounds for music signals constructed with plural instruments. We propose a method which detects an onset with a class separation using a deep neural network (DNN). To train the network, frames which include the waveform of the snare drum were extracted from music signal, and the log mel-filter bank channel outputs were calculated for each frame. The DNN was trained by using these feature parameters. The inputs were the calculated filter bank outputs, and the teaching signals were binary information whether to include the waveform of a snare drum or not. To detect onsets, candidate frames of the onset time were found based on changing in the power of the target music signal, and feature values were extracted from the candidate frames. Obtained values were used as the input data for the trained DNN, and detection of whether or not to include a waveform of a snare drum can be achieved. Experiments were conducted for 52 popular songs included in the music database, and an average F-measure was 0.73.

5pMU19. Computational estimation of released decades for Japanese women idol music from musical audio. Sota Okada (Dept. of Media Informatics, Graduate School of Ryukoku Univ., 1-5, Yokoe, Oe-Cho, Otsu, Shiga 520-2153, Japan, xkre3x@gmail.com) and Masanobu Miura (Dept. of Media Informatics, Ryukoku Univ., Otsu, Shiga, Japan)

This study realizes computational estimation of released decades for Japanese women idol (or diva) music from musical audio. Even though the released year of music is explicitly rebelled on the music, the feeling of decade, for example "1980's, or old-fashioned music" is completely subjective so it is thought as difficult for computers to evaluate, yet people may feel it when listening to music. Such feeling is thought to be derived from musical arrangement, instruments, and so forth, yet previous study clarified the possibility to estimate based on acoustic feature of music. If such estimation is realized, it could contribute to enhance conventional recommendation system of music by giving a query: "newly released but old-fashion music." This study conducted to establish explicit criteria for audio parameters by using logistic regression. Proposed system uses 25 types of acoustic parameters for each musical audio as training data set, and estimates single decades by simple logistic under a 10-fold cross validation scheme. As a result, estimation result of accuracy rate is 58.56% and kappa statistic is 0.48, being almost comparative to that of subjective evaluation. The estimation system is then implemented in a lap-top compute, which will be demonstrated at presentation.

5pMU20. Differences of characteristics of music singing between in usual Karaoke and in Hitokara. Junko Matsumoto (Nagano College of Nursing, 1694, akaho, Komagane, Nagano 399-4117, Japan, matsumoto@nagano-nurs.ac.jp)

Recently some of people in Japan go to Karaoke by oneself. This is called Hitokara (solo Karaoke). We investigated the differences of characteristics of music singing between in usual Karaoke and in Hitokara. Participants were 103 university students. They were asked to fulfill a questionnaire about characteristics of music singing both in usual Karaoke and in Hitkoara. As a result, about one third of participants had at least once experience of Hitokara singing, and they preferred both usual Karaoke singing and Hitokara singing. However, they preferred the music with characteristics of elation or lightness in usual Karaoke singing, in Hitokara singing they preferred the music with various characteristics. In Hitokara singing, people sing actively and they would be able to cope well with various moods by singing various music.

5pMU21. An acoustic and perceptual evaluation of four historic recordings of Chinese narrative singing. Shawn L. Nissen (Commun. Disord., Brigham Young Univ., 128 TLRB, Provo, UT 84602, shawn_nissen@byu.edu) and Francesca R. Lawson (Comparative Arts and Letters, Brigham Young Univ., Provo, UT)

This study is an acoustic and perceptual musicological evaluation of four recordings of vocal renditions from two schools of Chinese narrative singing that were differentiated by gender during the first half of the twentieth century. The analyses were based on four renditions of "At Break of Day" by Liu Baoquan (1869-1942) and Xiao Lanyun (1923-1992) of the "male school", as well as Luo Yusheng (1914-2002) and Lu Yiqin (1933-) of the "female school." The musicological analysis was based on transcriptions of the recordings, comparing melody, rhythm, ornamentation, and tonality. The acoustic analyses involved measures of central tendency and variability for duration and fundamental frequency (F0) using Praat acoustic analysis software. All four vocalists exhibited a similar mean F0, however differences in relative duration, F0 range, F0 variation, and F0 slope were found across singers at both the phrase and syllable level. Differences within schools of singing (male and female) and between instructor and student (i.e., Liu Baoquan and Xiao Lanyun) were also noted. These acoustic differences provide additional insights for ethnomusicologists that may be unattainable through traditional methods of perceptual analysis.

5pMU22. Effects of swing position in one measure on the feeling of musical rhythm: Examination using the rhythm dividing every beat. Shimpei Ikegami and Sumi Shigeno (College of Education, Psych. and Human Studies, Aoyama Gakuin Univ., 4-4-25 Shibuya, Shibuya-ku, Tokyo 150-8366, Japan, s-ikegami@ephs.aoyama.ac.jp)

Swing is an expression technique used in the performance of popular music. Swing causes the feeling of wanting to move the body in accordance with the musical rhythm. Swing is expressed by dividing one beat into two temporally unequal parts (e.g., 2:1). This study examined the effects of a swing's position in one measure on the feeling of musical rhythm. Every beat of the duple meter rhythm was divided into two parts. Four conditions of the swing's position were prepared: the "every-beat swing," the "odd-beat swing" (the swing's position being the first beat), the "even-beat swing" (the swing's position being the second beat), and the "no swing" conditions. This study had 114 participants listening to musical rhythms presented with Scheffé's paired-comparison method and rated the degree of "wanting to move the body" and "pleasantness." The results showed that the "every-beat swing" condition got the highest rating scores for a feeling of "wanting to move the body." We consider that the partial presence of swing, irrespective of the swing's position, might hardly allow listeners to attune to the rhythm of duple meter music.

5pMU23. An electromyographic study of left hand between experts and novices in violin playing. Satoshi Obata and Eriko Aiba (The Univ. of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan, s-obata@hi.is.uec.ac.jp)

We recently recorded string clamping force on the violin during simple (without vibrato) tone production by expert violinists. The results of our studies showed that the peak force exceeded 4.5 N at slow tempi, which decreased to 1.7 N at fast tempi. However, subjective assessment of playing effort indicated an opposite trend; the players felt that playing at faster tempi was more strenuous. We also measured muscular activity along with finger kinematics. We found that the mean left forearm muscle activity increased with tempo. The findings suggest that the speed of finger movement is more related to the arm and hand discomfort when playing a violin, as compared to clamping force. In the present study, we investigated the expert-novice difference in the nature of the string clamping force and related muscles in the left hand and arm while playing a violin. Information regarding the string clamping force and left hand muscle activity can help in teaching novices about the magnitude and timing of the appropriate string stabilizing action during sound production. Surface electromyography (EMG) and string clamping force were measured for this purpose.

5pMU24. Estimation of tempo, timing, and melody for piano practice support systems. Shota Asahi, Satoshi Tamara, Satoru Hayamizu (Eng., Gifu Univ., 1-1, Yanagido, Gifu-shi, Gifu, Gifu 501-1112, Japan, asahi@asr.info.gifu-u.ac.jp), and Yuko Sugiyama (Chubu Gakuin College, Seki, Japan)

In piano learning as voluntary practice, it is difficult especially for beginners to judge whether their musical performances are appropriate in terms of timing and melody. Therefore, we have studied piano learning support systems for beginners. In this study, we introduce a method for the system, to estimate tempo, timing and melody of performances, and to visualize the data for piano learners and instructors. In our method, a tempo for each musical bar is firstly estimated by measuring first notes in current and incoming bars. Subsequently, a timing of each musical note in the bar is obtained based on the estimated tempo. A chroma vector is then computed for each note, in order to compare it with the vector obtained from a corresponding instructor's performance, and finally to detect misplayed keys. In order to confirm effectiveness of our method, we conducted experiments. We collected musical performances (No.12 Bayer) from seven subjects. Assuming voluntary practice, the subjects were asked to listen to music and watch corresponding movies in which an instructor played piano with musical scores and instructions, before they made performances. We analyzed the performance data and conducted questionnaire to the subjects, to confirm the effectiveness.

5pMU25. Mouthpiece pressing force for pitch and loudness control in playing the French horn. Takeshi Hirano (J. F. Oberlin Univ., 3758 Tokiwa-machi, Machida-shi, Tokyo 194-0294, Japan, hira.take.713@gmail.com) and Hiroshi Kinoshita (Osaka Aoyama Univ., Osaka, Japan)

Using a force-sensing French horn mouthpiece developed by the present authors, lip pressing force was examined during pre-attack and post-attack phases of 2-sec sustained tone production at varied pitches (87, 174, 349, 466, and 698 Hz) and dynamics (95, 100, 105, 110, and 115 dB) performed by 12 highly trained horn players. Dependent variables examined were: force at the onset of sound, a peak of the force rate before the sound onset, and mean and SD forces during steady state tone production. All of these force variables increased nearly linearly with pitch. They also increased lightly with louder dynamics. The pitch-related and dynamic-related increase of the force varied largely among the players. These findings indicate that skilled horn performances are supported by pre-attacking the mouthpiece pressure properly to prepare for a target pitch at target dynamics. The notable inter-player difference in force production in skilled players suggests a proper level of mouthpiece pressures witch differs among players due possibly to congenital physical properties of their lips.

5pMU26. Spatial acoustic-radiation characteristic of a violin captured in the vicinity of a violinist. Katuhiro Maki (Faculty of Human Informatics, Aichi Syukutoku Univ., 2-9 Katahira, Nagakute, Aichi 480-1197, Japan, maki-ns@umin.ac.jp)

The spatial characteristic of the sound of a violin is thought to affect the timbre of a violin in a reverberant space such as a concert hall. Therefore, the spatial characteristic is one of important key to clarify the timber of the violin. In this study, the spatial characteristic was investigated by placing 42 small microphones around a violinist and recording the sound of the violin being played. The experiment—targeting nine modern violins—was performed in an anechoic chamber. The professional violinists were asked to perform musical scales and compositions. According to the results of an analysis of the sound, in the horizontal plane, the sound is radiated roughly toward diagonally forward left of the violinist; in the sagittal plane, sounds with frequency components lower than around 1 kHz are radiated downwards, and sounds with frequency components higher than 1 kHz are radiated upwards. Moreover, it was observed that the directionality of the radiated sound of the violin is particularly strong around 1 kHz and 3 kHz, and somewhat strong near 350 Hz. It was also observed that during the performances of the compositions, the direction of the sound radiation at each analyzed frequency varies significantly with time.

Session 5pNSa

Noise: Design of Noise Control Materials

J. S. Bolton, Cochair
Ray W. Herrick Laboratories, School of Mechanical Engineering, Purdue University, Ray W. Herrick Laboratories, 177 S. Russell St., West Lafayette, IN 47907-2099

Takashi Yamamoto, Cochair
Mechanical Engineering, Kogakuin University, 2665-1 Nakano, Tokyo 1920015, Japan

Chair's Introduction—1:00

Invited Papers

1:05

5pNSa1. A level set-based topology optimization for fluid-structure coupled problems by using two-phase material model. Takashi Yamamoto (Mech. Eng., Kogakuin Univ., 2665-1 Nakano, Tokyo, Hachioji 1920015, Japan, takashi_yamamoto@cc.kogakuin.ac.jp), Yuki Noguchi, Takayuki Yamada, Kazuhiro Izui, and Shinji Nishiwaki (Graduate School of Eng., Kyoto Univ., Kyoto, Japan)

In this study, we develop a two-phase material model mixed with a solid and fluid phases to represent an elastic material and an acoustic material in a unified set of governing equations. By applying the model, boundary conditions at boundaries between different two-phase materials are satisfied naturally, and the boundaries do not have to be expressed explicitly. This characteristic is effective for topology optimization of fluid-structure coupled problems by using finite element method. Consequently, we formulate an optimization method utilizing the two-phase material model and a level set method to express boundaries between different two-phase materials. An approximate method to obtain topological derivative is also developed by applying the variational analysis and the adjoint variable method. Topological derivative of an objective functional is considered to be proportional to the derivative of the level set functions with respect to time and the values of the level set functions are updated by solving time evolutionary reaction diffusion equation with a regularization term. After the verifications of the two-phase model, we present several numerical demonstrations for directional acoustic cloaking of a cylindrical obstacle placed in unbounded air medium.

1:25

5pNSa2. Level set-based topology optimization for the design of acoustic metamaterial using two-phase material model. Yuki Noguchi (Dept. of Mech. Eng. and Sci., Kyoto Univ., Kyoto daigaku-katsura C3, Nishikyo-ku, Kyoto 615-8540, Japan, noguchi.yuuki.34s@st.kyoto-u.ac.jp), Takashi Yamamoto (Dept. of Mech. Eng., Kogakuin Univ., Tokyo, Japan), Takayuki Yamada, Kazuhiro Izui, and Shinji Nishiwaki (Dept. of Mech. Eng. and Sci., Kyoto Univ., Kyoto, Japan)

Acoustic metamaterials are artificial materials composed of periodic structures made of elastic and fluid material. They could be used for noise reduction problems and efficient sound propagation control by their unusual properties, such as negative bulk modulus, mass density, and refractive index. The performance of acoustic metamaterial is strongly affected by its designs rather than its material properties. Therefore, topology optimization method could be used to obtain a desired property of acoustic metamaterial. To model acoustic-elastic coupled system, we introduce a two-phase material model where the solid and fluid phases are mixed. While coupling boundary conditions are required in a usual acoustic-elastic coupled analysis, this model allows us to analyze the system without imposing such boundary conditions since their conditions are satisfied naturally. In this research, we propose a level set-based topology optimization method for the design of acoustic metamaterial. First, we introduce the two-phase material model and level set-based topology optimization method. An optimization problem for an acoustic metamaterial that has negative mass density is formulated and sensitivity analysis is performed based on the concept of topological derivative. After verifying the topological derivative by comparing it to numerical differences, we present two-dimensional numerical examples for negative mass acoustic metamaterial.

1:45

5pNSa3. Acoustic media with pore size distribution and assignment of the pore characteristic lengths and permeabilities. Kirill V. Horoshenkov (Univ. of Sheffield, Mappin St., Sheffield S1 3JD, United Kingdom, k.horoshenkov@sheffield.ac.uk), Matti Niskanen, Jean-Philippe Groby, Olivier Dazel, and Aroune Duclos (de l'Universitée du Maine, de l'Universitée du Maine, Le Mans, France)

The model by Champoux and Allard [J. Appl. Phys. **70**(4), 1975-1979 (1991)] is used extensively for research and development of new porous media solutions and for predicting sound propagation in the presence of porous media. Four key non-acoustical parameters in this model which are rarely measured non-acoustically are: the viscous and thermal characteristic lengths, thermal permeability and Pride parameter Pride *et al* [Phys. Rev. B **47**, 4964-4978 (1993)]. In this work we show how these parameters unambiguously relate to the pore size distribution which is a characteristic measured nonacoustically and more routinely. We compare the predictions by this model against the log-normal pore size distribution model by Horoshenkov *et al.* [J. Acoust. Soc. Am. **104**, 1198-1209 (1998)] and show that these two models provide very close predictions when the four non-acoustic parameters are expressed through the mean pore size

and its standard deviation. We also confirm these results through a 4-microphone impedance tube experiment in which we determine the dynamic density and complex bulk modulus of a range of porous media. We also study the values of the viscous and thermal characteristic lengths and viscous and thermal permeabilities which fit our acoustic data.

2:05

5pNSa4. Effect of membranes at microscopic polygonal faces of polyurethane foam by using homogenization method. Takashi Yamamoto and Yuki Imae (Mech. Eng., Kogakuin Univ., 2665-1 Nakano, Hachioji, Tokyo 1920015, Japan, takashi_yamamoto@cc. kogakuin.ac.jp)

Sound absorption coefficient in poroelastic media depends on the geometries in the microscopic scale. In some polyurethane foams that exhibit high sound absorption coefficients in the middle frequency range around 1 kHz, the faces of polygonal shape of microscopic structure are covered by thin membranes. Moreover, a hole is frequently found in some membranes. This type of microscopic structure is considered to contribute to the high sound absorption performance. Multi-scale analysis for sound-absorbing poroelastic media with periodic microscopic geometries has been recently proposed by one of the authors. In this method, the homogenization method based on the asymptotic expansions are utilized and further extended to express the viscous dissipation in the vicinity of the boundary between the solid and fluid phases, thermal dissipation from the fluid phase to the solid phase. Analysis in the microscopic scale is first performed by using a unit cell of the periodic structure and macroscopic properties are derived. The properties are then applied to calculate macroscopic response such as sound absorption coefficient. In this paper, we apply this multi-scale analysis and investigate numerically the relationship between the size of the hole in membranes and sound absorption performance.

Contributed Papers

2:25

5pNSa5. Design and optimization of membrane-type acoustic metamaterials using genetic algorithms. Matthew G. Blevins (Durham School of Architectural Eng. and Construction, Univ. of Nebraska - Lincoln, 408 Ginger Bend Dr. Apt. 207, Champaign, IL 61822, matthew.g.blevins@usace.army.mil), Siu-Kit Lau (Dept. of Architecture, National Univ. of Singapore, Singapore, Singapore), and Lily M. Wang (Durham School of Architectural Eng. and Construction, Univ. of Nebraska - Lincoln, Omaha, NE)

One of the most difficult problems in noise control is the attenuation of low frequency noise. Typical solutions require barriers with high density and/or thickness. Membrane-type acoustic metamaterials are a novel type of engineered material capable of high low-frequency transmission loss despite their small thickness and light weight. These materials are ideally suited to applications with strict size and weight limitations such as aircraft, automobiles, and buildings. This paper reviews computationally efficient dynamic models for designing membrane-type acoustic metamaterials that have been proposed by the authors, based on the impedance-mobility approach. The computational efficiency of the impedance-mobility models compared to finite element methods enables implementation in design tools based on a graphical user interface and in optimization schemes. Genetic algorithms are used to optimize the unit cell design for a variety of noise reduction goals, including maximizing transmission loss for broadband, narrow-band, and tonal noise sources. The degree to which the optimal configurations extend to systems of multiple unit cells and unit cells of different shapes is explored.

2:40

5pNSa6. An application of a parametric transducer to measure the acoustical properties of a living green wall. Anna Romanova (Eng. & Sci., Univ. of Greenwich, Central Ave., Chatham Maritime, Chatham, Kent ME4 4TB, United Kingdom, a.romashk@gmail.com) and Kirill V. Horoshenkov (Mech. Eng., Univ. of Sheffield, Sheffield, United Kingdom)

Greening of urban spaces provides a number of environmental benefits. Living Green Walls (LGW) is a most typical example of greening which is also known for its ability to absorb unwanted noise. However, this ability of LGW to absorb noise is rather hard to quantify, because there is a lack of reliable experimental methods to measure it in-situ. This work reports on a new method to measure the absorption coefficient of LGW which makes use of a highly directional parametric transducer and acoustic intensity method. This method is tested under controlled laboratory conditions and in a typical street environment. The results of these experiments demonstrate the ability of the method to measure the absorption of a LGW. It also enables us to quantify the effects of the plant type and moisture content in the soil on the ability of the LGW to absorb sound. The proposed method has certain benefits over ISO354-2003 and CEN/TS 1793-5:2003 standard methods.

2:55

5pNSa7. Adaptive algorithms for combined feedforward-feedback controllers. Jacob Bean, Chris Fuller (Aerosp. Eng., Virginia Tech, 100 Exploration Way, Hampton, VA 23666, bjacob09@vt.edu), and Noah Schiller (Structural Acoust. Branch, NASA Langley Res. Ctr., Hampton, VA)

Traditional approaches to feedback and hybrid active noise and vibration control are generally based on adaptive gradient search methods using estimates of the instantaneous gradient. Most of these algorithms employ internal model control and are derivatives of the least mean squares algorithm commonly used in adaptive filtering. In these algorithms, the gradient calculation neglects the presence of a feedback path. To avoid potential algorithm divergence due to the presence of the feedback path, it is common to update the filter conservatively by including a leakage term, which is in general a trial and error procedure. This paper presents mathematical derivations for the updating mechanisms in feedback and hybrid controllers that make no assumptions regarding the accuracy of the plant model. Rather than deriving update equations assuming an ideal plant model and then modifying the adaptation process to be less sensitive to model errors, the reverse approach is taken. It will be seen that the presence of the feedback loop can be accounted for directly in the update equation. This allows either hard or probabilistic bounds on the uncertainty to be set at the outset of design. Computer simulations are presented to validate the algorithms.

3:10

5pNSa8. Characteristic analysis of screeching noise using 24bit/192 kHz high-resolution measurement system in anechoic room. Shuya Ogino and Kan Okubo (Graduate School of System Design, Tokyo Metropolitan Univ., Asahigaoka 6-6, Hinosi, Tokyo 191-0065, Japan, shuwashuwa555@gmail.com)

Many researches on the screeching sound have been investigated for some decades. Halpern et al.(1986) reported acoustic evaluation of "chilling sound" as a pioneering study. The screeching sound makes us feel so-called hair-raising, bloodcurdling chills, which is generated by scratching a blackboard and frosted glass with a fingernail, metal, and so on. This might be common sensation shared by the world. Although Halpern et al. (1986) employed a tape recorder of up to 20.0 kHz as a recording equipment, in recent years researchers can use a 24 bit/192 kHz high-resolution measurement system. The frequency range of screeching sound can't be not fundamentally limited up to 20.0 kHz. Scratching should cause the sound above the human audible range; it is known as "Hypersonic effect." In this study, therefore, we try to record and analyze the screeching noise at the frequency range of up to 96 kHz. We employ 24 bit/192 kHz high-resolution measurement system in anechoic room as a recording equipment. This presentation reports the findings from characteristic analysis of screeching sound obtained by high-resolution measurement. We demonstrate that scratching generates the sound above 20.0 kHz, and that so-called screeching noise have many overtone sounds above 20.0 kHz and remarkable features of like sawtooth wave.

Session 5pNSb

Noise: Emerging Issues on Quiet Vehicles—A Matter of Soundscape

Brigitte Schulte-Fortkamp, Cochair
Institute of Fluid Mechanics and Engineering Acoustics, TU Berlin, Einsteinufer 25, Berlin 101789, Germany

Klaus Genuit, Cochair
HEAD acoustics GmbH, Ebertstr. 30a, Herzogenrath 52134, Germany

Katsuya Yamauchi, Cochair
Faculty of Design, Kyushu University, Shiobaru 4-9-1, Fukuoka 815-8540, Japan

Chair's Introduction—3:40

Invited Papers

3:45

5pNSb1. E-vehicle: A positive impact on Soundscape? Klaus Genuit (HEAD Acoust. GmbH, Ebertstr. 30a, Herzogenrath, NRW 52134, Germany, Klaus.Genuit@head-acoustics.de)

The future will see more and more electric cars on the road as stronger exhaust emission regulations are put into effect. A byproduct of this shift is a broad reduction of noise: no combustion-, exhaust- or intake-noise in addition to reduced tire-road noise. On the other hand, concerns have been raised with respect to pedestrian safety and electric vehicles being too quiet. The significant impact on the soundscape manifested by electric vehicles will be described in this paper, as well as efforts to address safety concerns. The request to compensate the low noise level of electrical vehicles by acoustical warning signals should be discussed very carefully, as even low level warning signals could increase perceived annoyance. This scenario will be explained by several examples.

4:05

5pNSb2. Relationship between the use of portable audio devices and taking notice of the approach informing sound of quiet vehicles. Mariko Hamamura, Shotaro Takamaki, and Yasunari Obuchi (School of Media Sci., Tokyo Univ. of Technol., 1404-1 Katakuracho, Hachioji, Tokyo 192-0982, Japan, mariko@hamamura.biz)

Lively discussion about the approach informing sound of quiet vehicles is being conducted; however, people often listen to media on portable audio devices in various environments and may not notice the approach informing sound. To clarify the relationship between the use of portable audio devices and the experience of encountering dangers associated with the approach of quiet vehicles, a questionnaire survey was conducted with 158 university students. Portable audio device ownership was 99.3%; only one respondent reported not owning such a device. Twenty-two respondents (15.5%) had encountered dangerous situations while using these devices outdoors. Thirteen respondents reported that they had experienced near-misses with vehicles while walking or biking, and seven reported that the vehicles were quiet vehicles. Multiple correspondence analysis revealed some differences in the usage habits of portable audio devices between the users who had and had not encountered dangerous situations. To improve notification effectiveness of quiet vehicles, usage habits related to portable audio devices, especially concerning young people, should be considered. In addition, means of preventing portable audio device users' tendency to "tune out" other sounds should be addressed.

4:25

5pNSb3. Computational analysis of aeroacoustic noise generated from a rotating tire with a longitudinal groove. Daisuke Kato, Gonghwi Lee (Aeronautics and Astronautics, Univ. of Tokyo, Yoshinodai,Sgamihara,Kanagawa,Japan, Sagamihara, Kanagawa 2520206, Japan, dkato@flab.isas.jaxa.jp), Yoshiaki Abe (Tokyo Inst. of Technol., Yokohama, Japan), Taku Nonomura, Akira Oyama (Japan Aerosp. Exploration Agency Inst. of Space and Astronautical Sci., Sagamihara, Kanagawa, Japan), Kozo Fujii (Tokyo Univ. of Sci., Katsushika, Japan), Toshiyuki Ikeda, and Masataka Koishi (The Yokohama Rubber, Hiratsuka, Japan)

It is known that acoustic waves which is higher than 1000 Hz are generated from a rotating tire, and one possibility of their generation mechanism is aeroacoustic effect. The main objective of this study is to understand the generation mechanism and acoustic wave characteristic for realization of low noise tire. For this main objective, the effects of a longitudinal groove shape on acoustic field are investigated in the present study. Three cases, without a groove case (slick tire), a small longitudinal groove and large longitudinal groove are analyzed by large eddy simulations. The result shows that compression or expansion of the vortex is alleviated and weaken sound source by attaching a groove on tire surface. On the other hand, we also confirm that a groove contributes generation of new fluctuation and affects the spread of fluctuations. It can be considered that acoustic fields of front side of tire are affected by a tradeoff between the effect of compression or expansion of the vortex and those of the spread of fluctuation.

5pNSb4. Effectiveness of additional warning sounds for quiet vehicles in urban noise environment. Katsuya Yamauchi (Faculty of Design, Kyushu Univ., Shiobaru 4-9-1, Fukuoka 815-8540, Japan, yamauchi@design.kyushu-u.ac.jp)

Relative quietness of electric or hybrid electric vehicles is one of the most important topics in environmental acoustics. Although reduced vehicle noise is eligible for urban sound environment, it is also a matter of pedestrians' safety concern. Hence, regulations regarding additional warning sounds for the quiet vehicles have been developing in Japan as well as in global. Recently the UN-regulation regarding the warning system to compensate their quietness by additional warning device, Acoustic Vehicle Alerting System (AVAS) according to the regulation, has been approved. Those kind of measures should be carefully discussed. Several studies have been conducted in different institutions to examine the feasible sound design for the warning to be detected in urban noise environment. This paper provides a cross-cutting perspective on the detectability of the AVAS sound compared to the background noise levels, which illustrates that the AVAS could solve problems only in limited scenarios.

5pNSb5. The move toward quiet vehicles—A paradigm shift. Brigitte Schulte-Fortkamp (Inst. of Fluid Mech. and Eng. Acoust., TU Berlin, Einsteinufer 25, Berlin 101789, Germany, b.schulte-fortkamp@tu-berlin.de)

The introduction of so-called quiet vehicles is not simply a matter of acoustics. It also presents a major paradigm shift in both the car industry and daily life as a whole. A remaining consideration is how adequate public communication can support the acceptance of quiet vehicles. Soundscape is a construct of human perception, influenced by one's socio-cultural background as well as the acoustic environment in context. Among many concerns, the meanings of sound, the composition of diverse sound sources, the listener's attitude, and expectations towards the acoustic environment are particularly important in one's perception of a soundscape. Previous experiences with internal-combustion engines in vehicles significantly inform people's different perceptions and assessments of the environment, and likewise are important factors when considering the introduction of quiet cars in society. The discussion in this paper will show how the soundscape approach can be used to involve people in assessing the needs and uses of these "new cars."

5:25–5:45 Panel Discussion

FRIDAY AFTERNOON, 2 DECEMBER 2016   SOUTH PACIFIC 2, 1:30 P.M. TO 4:00 P.M.

Session 5pPA

Physical Acoustics: Radiation Force and Nonlinear Acoustics

Kyle S. Spratt, Chair
Applied Research Laboratories, The University of Texas at Austin, Austin, TX 78713

Contributed Papers

1:30

5pPA1. Acoustic radiation force on non-spherical objects in a liquid. Bart Lipkens (FloDesign Sonics, 1215 Wilbraham Rd., Box S-5024, Springfield, MA 01119, blipkens@wne.edu), Yurii A. Ilinskii, and Evgenia A. Zabolotskaya (FloDesign Sonics, Austin, Texas)

Macro-scale acoustophoretic separation has applications in oil/water separation and biopharmaceutical cell clarification. Interesting phenomena include deformation of oil droplets and clustering of cells into cylindrical shapes. Particle and cluster size are not necessarily small relative to the wavelength. A general computational framework is needed to calculate the acoustic radiation force on particles, clusters, and cells of various shapes. Ilinskii *et al.* presented a method to evaluate an acoustic radiation force on spherical objects, based on expansion of incident and scattered field in spherical waves. The acoustic force is given by the incident wave amplitudes and dimensionless scattering coefficients. This method is generalized for particle shapes which are not spherical. The incident acoustic and scattered field are expanded with respect to spherical waves. An evaluation of the scattering amplitudes for non-spherical objects can be done numerically. Analytically, scattering amplitudes are calculated from boundary conditions for spheroidal functions that are solutions of a wave equation in spheroidal

coordinates ξ, η, φ. Boundary conditions are defined at surface ξ= constant, i.e. on an ellipsoid surface. Solutions for prolate and oblate ellipsoids have been estimated. The asymptotic values of spheroidal solutions are determined for ξ large enough to obtain required amplitudes of scattered field. Ilinskii et al. POMA, **19** (2013).

1:45

5pPA2. Negative pulling acoustic radiation force on spheroids in Bessel vortex beams. Farid G. Mitri (Area 52 Technol., Chevron, 5 Bisbee Court, Santa Fe, NM 87508, f.g.mitri@ieee.org)

The emergence of a negative pulling acoustic radiation force on rigid spheroids is examined, with the use of acoustical Bessel vortex (helicoidal) "*tractor*" beams. Numerical predictions based on the acoustical scattering phenomenon [F. G. Mitri, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **62**, 1809-1818 (2015); Ann. Phys. (NY) **363**, 262-274 (2015)] that is extended to calculate the longitudinal acoustic radiation force (i.e., acting along the direction of wave propagation) illustrate the theory, with particular emphasis on the aspect ratio of the spheroid, the half-cone angle and order of the beam, as well as the dimensionless frequency parameter. It is shown that the Bessel vortex beam composed of progressive (i.e., traveling) waves

is capable of generating a negative pulling force acting in opposite direction of the forward linear momentum density flux associated with the incident waves [F. G. Mitri, Europhys. Lett. **112**, 34002 (2015)]. The scope of the analysis is further expanded to include the cases of standing and quasi-standing Bessel beam tweezers [F. G. Mitri, Wave Motion **57**, 231-238 (2015)]. The results potentially suggest the use of Bessel "*tractor*" vortex beams for pulling, propelling, or manipulating arbitrary-shaped objects deviating from the spherical geometry for applications in remote sampling, particle manipulation, and handling, to name a few areas.

2:00

5pPA3. Study on the non-contact trapping of a small object using a pipe radiating travelling waves and a reflective rod—The effect of the shape of reflector surface on the trapping state. Manabu Aoyagi (Muroran Inst. of Technol., 27-1, Mizumoto, Muroran, Hokkaido 050-8585, Japan, maoyagi@mmm.muroran-it.ac.jp), Hideki Tamura, and Takehiro Takano (Tohoku Inst. of Technol., Sendai, Japan)

Flexural damped travelling waves can be excited in a lossy transmission pipe. The waves generate travelling acoustic flow in the inside of pipe and near the pipe end and the acoustic flow serves as travelling acoustic waves. It has been confirmed that a non-contact transportation of a small object can be performed in the pipe, and additionally, a small object can be trapped at a nodal point of the standing wave formed by placing a reflector rod near the pipe end. In the experiment, a polycarbonate pipe (Outer dia.:13 mm, Inner dia.:9 mm, Length:160 mm) for the lossy ultrasonic transmission pipe that a piezo-ceramic annular plate vibrating in the first radial vibration mode is adhered at the pipe end was tried and the study on the effect of the shape of reflector surface on the trapping state is considered. The reflector is made of an acrylic rod (Dia.:13 mm) and the reflective surface is formed in to the conical shape. It was confirmed that the conical shape to be adapted for the trapping exists. To improve in the trapping force of the device, we considered the pipe attaching a reflective ring at the pipe end. The apparatus and distinctive characteristics are shown.

2:15

5pPA4. Trapping and manipulation of a small object in the space between two pipes placed to the counter direction. Takehiro Takano, Hideki Tamura (Tohoku Inst. of Technol., 35-1 Kasumi-cho, Yagiyama, Sendai 982-8577, Japan, ktakano@tohtech.ac.jp), and Manabu Aoyagi (Muroran Inst. of Technol., Muroran, Japan)

A small object can be trapped in the space between two pipes which are placed to the counter direction and radiated traveling acoustic waves from the pipe ends. In the study, two pipes vibrating in axisymmetric flexural damped traveling wave to generate acoustic flow in the pipe inside are used. The flow propagates in the pipe inside and is radiated for traveling acoustic waves from the pipe end. The pipe made of polycarbonate is used for ultrasonic transmission pipe and two piezo-ceramic annular plates are adhered at the each pipe end for exciting the vibration modes. The first radial vibration mode of the annular plate is used and its resonance frequency is about 59 kHz. When the two vibrators are driven in opposite phase, a small object can be trapped at intermediate point in the space between two pipe ends. The gap of the space is about 10 mm. The trapping object can be shifted its position corresponding to the phase difference of the two driving voltages. That is, the object can be moved on non-contact manipulation in the space between the two pipe ends. These results are confirmed experimentally and by simulation according to FEA.

2:30

5pPA5. Non contact pick-up of drop of liquid from micro-well with megahertz ultrasound. Hiroki Tanaka, Yosuke Mizuno, and Kentaro Nakamura (Lab. for Future Interdisciplinary Res. of Sci. and Technol., Tokyo Inst. of Technol., R2-26, 4259 Nagatsuta, Yokohama, Kanagawa 226-8503, Japan, htanaka@sonic.pi.titech.ac.jp)

In recent years, considerable attention has been paid to the noncontact transport of small objects and droplets as an important technology in pharmacy industries, and new material science/engineering. Our aim in this report is to develop a noncontact pipette, and we focused on the method to pick-up a single droplet from micro-well plate without contact. A focused immersion transducer is placed below a micro-well plate. By emitting burst waves at 1.5 MHz, a droplet is ejected from the well and caught by the pressure nodal plane in the standing wave field. First, by observing the deformation of the ejected droplet and the surface of liquid in the micro-well plate with a high-speed video camera, we discussed the principle of droplet pick-up. Next, in order to pick-up any type liquid, we investigated the effect of surface tension and viscosity on the noncontact ejection of droplet for water, ethanol, and silicone oil. In this experiment, we found a linear relationship between the ejected droplet volume and the surface tension. In addition, we discuss the ratio of Ohnesorge number to Reynolds number in association with the stability of drop-on-demand inkjet printing operations, and determine the condition for droplet jetting.

2:45–3:00 Break

3:00

5pPA6. Shape and rotation analysis on an ultrasonically levitated droplet using distributed point source method and least square moving particle semi-implicit method. Yuji Wada, Kohei Yuge (Faculty of Sci. and Technol., Seikei Univ., 3-3-1 Kichijoji-kitamachi, Musashino, Tokyo 180-8633, Japan, yuji.wada@st.seikei.ac.jp), Hiroki Tanaka, and Kentaro Nakamura (Lab. for Future Interdisciplinary Res. of Sci. and Technol., Tokyo Inst. of Technol., Yokohama, Kanagawa, Japan)

Ultrasonic droplet levitation has recently been drawing attention as a way of non-contact transportation. Many experiment revealed that the shape changes or streaming field on the levitating droplets, however, quite small number of numerical simulation have discussed this phenomenon including fluid dynamics within the droplet. In this paper, coupled analysis using the distributed point source method (DPSM) and the least square moving particle semi-implicit (LS-MPS) method, both of which do not require grids or meshes to handle the moving boundary with ease, is suggested. The acoustic radiation force including viscoacoustic torque, which emerges from the viscous boundary layer, is calculated from the distributed point source method result using the idea of boundary layer normal velocity and input to the LS-MPS surface particles. A droplet levitated in an acoustic chamber is simulated using the proposed calculation method. The droplet is vertically supported by a plane standing wave from an ultrasonic driver and subjected to a rotating sound field excited by two acoustic sources on the side wall with different phases. The rotation of the droplet is successfully reproduced numerically and its acceleration is discussed and compared with those in the literature.

3:15

5pPA7. Pressure dependence of the attenuation and sound speed in a bubbly medium: Theory and experiment. Amin Jafari Sojahrood (Dept. of Phys., Ryerson Univ., 350 Victoria St., Toronto, ON M5B2K3, Canada, amin.jafarisojahrood@ryerson.ca), Qian Li (Biomedical Eng., Boston Univ., Boston, MA), Hossein Haghi, Raffi Karshafian (Dept. of Phys., Ryerson Univ., Toronto, ON, Canada), Tyrone M. Porter (Biomedical Eng., Boston Univ., Boston, MA), and Michael C. Kolios (Dept. of Phys., Ryerson Univ., Toronto, ON, Canada)

Existence of bubbles in a medium changes the attenuation and sound speed (Cs) of the medium. The relationship between the nonlinear oscillations of the MBs and acoustical parameters of the medium is not fully understood. In this work, monodisperse solutions of lipid coated bubbles with mean diameter of 5.2 micron and peak concentration of (5000 microbubbles/ml) were sonicated with a broadband pulse with center frequency of 2.25 MHz. The attenuation and Cs, were measured over a pressure range of 10-100 kPa. Using our recent nonlinear model and by solving the Marmattont Model, the attenuation and Cs of the solution were numerically analyzed. Experimental results showed that as the pressure increased, the attenuation peak increased by ~7-8 dB while its frequency decreased from 2.05 to 1.55 MHz. The maximum Cs of the medium increased with pressure (1521 to 1529 m/s) and shifted towards lower frequencies (2.34 to 1.95 MHz). At a fixed frequency (e.g., 1.65 MHz) the Cs increased by ~14%. The Cs of the bubbly medium remained the same as the Cs in the absence of bubbles at the frequency of the attenuation peaks. Numerical simulations were in good agreement with experimental observations and confirm the experimentally observed phenomenon.

5pPA8. Evaluation of nonlinear effects at the focus of a hemispherical acoustic source. Oleg A. Sapozhnikov, Pavel B. Rosnitskiy, and Vera A. Khokhlova (Phys. Faculty, Moscow State Univ., Leninskie Gory, Moscow 119991, Russian Federation, oleg@acs366.phys.msu.ru)

Several theoretical models exist for simulating nonlinear generation of harmonics for weakly or moderately focused waves. Under these conditions, the acoustic field can be considered as a directed beam. In the case of hemispherical sources, this assumption is no longer valid, so development of different models is needed. Here, a theoretical approach is presented for evaluating the amplitude level of the second harmonic generated in hemispherically focused fields. The approach is based on the successive approximation method and includes several steps. First, the linear acoustic problem is solved for a hemispherical source using a spherical harmonics series expansion. Next, the volume sources for the second harmonic generation are calculated using the solution to the linear problem. Finally, these sources are used for obtaining an analytic solution for the second harmonic amplitude at the focus. The developed method was applied to study the acoustic field of the ExAblate 3000 clinical transducer that has a hemispherical shape of 30 cm diameter, operating frequency of 670 kHz, and acoustic power of several hundred watts. Calculations predict that nonlinear effects are weak even for the free-field focusing in water at acoustic power levels up to 1 kW. [Work supported by RSF 14-15-00665.]

5pPA9. Introducing Hugonions: Particle-like shocks that travel, interact, and annihilate. Jae-Wan Lee and Won-Suk Ohm (School of Mech. Eng., Yonsei Univ., Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, South Korea, fantalex@yonsei.ac.kr)

This talk presents a novel computational method for 1-D wave processes in fluids, the applicability of which extends to strongly-nonlinear waves of arbitrary strength. A common approach to modeling waves of arbitrary strength in fluid dynamics is the Riemann solver of the Euler equations, requiring enormous computation time and resources. In this study, we expand upon our previous work [Lee et al., AIP Conf. Proc. **1685**, 070011 (2015)], in which the propagation of a strongly-nonlinear wave is described by a collection of forward-traveling, particle-like shocks, dubbed "Hugonions" for their adherence to the Rankine-Hugoniot relations. To extend the applicability of Hugonions to compound waves, the Riemann solution of an arbitrary discontinuity is augmented to the defining characteristics of Hugonions. This way, Hugonions behave like particles, which travel, interact with each other, and annihilate if certain conditions are met. Because a computational mesh is not necessary (i.e., the method simply keeps track of Hugonions), and most importantly Hugonions will reduce in numbers following successive interactions, our computational method leads to a few orders-of-magnitude reduction in computation time compared to the existing Riemann solver-based schemes. Verification of the method against the Riemann solver is performed for a number of well-known benchmark problems.

Session 5pSCa

Speech Communication: Speech Processing

Matthew Goupell, Chair
Hearing and Speech Sciences, University of Maryland, College Park, 0119E Lefrak Hall, MD 20742

Contributed Papers

1:00

5pSCa1. Better-ear glimpsing in Japanese adult bilateral cochlear implant users. Baljeet Rana, Jörg M. Buchholz (National Acoust. Labs. (NAL), Level 5, Australian Hearing Hub 16 University Ave., Sydney, NSW 2109, Australia, baljeet.rana@nal.gov.au), Catherine Morgan (Cochlear Asia pacific, Sydney, NSW, Australia), Mridula Sharma (Dept. of Linguist (Audiol. Section), Macquarie Univ., Sydney, NSW, Australia), Tobias Weller (National Acoust. Labs. (NAL), Sydney, NSW, Australia), Shivali A. Konganda (Dept. of Linguist (Audiol. Section), Macquarie Univ., Sydney, NSW, Australia), Kyoko Shirai, and Atsushi Kawano (Dept. of Otolaryngol., Tokyo Medical Univ., Tokyo, Japan)

Better-ear glimpsing (BEG) is a phenomenon that helps understanding speech in the presence of fluctuating, spatially separated distractors. This phenomenon has been studied in normal-hearing (NH) and hearing-impaired listeners but remains untapped in adults with cochlear implants (CIs). Further, it has not been investigated how far providing CIs in both ears can improve BEG over providing a single CI in one ear. In the current study, seven Japanese adult bilateral CI users with post-lingual deafness were recruited. Male and female speech distractors were presented from $+ 90^0$ and—90^0 and target questions taken from the Helen sentence test (translated into Japanese, spoken by a native Japanese female speaker) were presented from the front. Speech comprehension thresholds (SCTs) were measured in both a co-located and spatially separated condition with CIs in both ears and only in one ear. BEG was calculated as the difference in SCTs between the co-located and spatially separated condition. BEG noted with bilateral CIs was about 2 dB and with unilateral CI was about 0.3 dB, which is significantly smaller than found in NH listeners. The bilateral benefit noted was about 3 dB indicating the advantage of having CIs in both the ears.

1:15

5pSCa2. Challenges in understanding and modeling speech perception. Adelbert W. Bronkhorst (Human Factors, TNO, Kampweg 5, Soesterberg 3769DE, Netherlands, adelbert.bronkhorst@tno.nl)

Speech perception performance is one of the most widely used measures in audiology and auditory research. Nevertheless, our understanding of how speech is processed by the auditory system is limited. This is not only due to the subject—the complexity of the speech signal itself and the intricacy of disentangling target speech from interfering sounds—but also to the way it is approached: different aspects of speech perception are studied by largely separate research communities. Also when only early processing of speech is considered, excluding psycholinguistic aspects and effects of memory and learning, at least three relevant research fields can be discriminated: audibility and masking, auditory scene analysis, and attention. Although there is increasing interest in the crosslinks between these fields, approaches are still insufficiently integrated and important questions remain unanswered. Furthermore, modeling efforts are lagging behind and are only partly driven by advances in neurophysiological research. After reviewing progress in the three fields, a conceptual model will be presented that attempts to tie results together. Finally, knowledge gaps and open research issues will be highlighted.

1:30

5pSCa3. Unusual auditory/speech processing in autism spectrum disorder revealed by verbal transformation effects. Chihiro Itoi (Chuo Univ., 742-1 Higashinakano Hachioji-shi, Tokyo 192-0393, Japan, itoi.chihiro@gmail.com), Nobumasa Kato (Medical Inst. of Developmental Disabilities Res., Showa Univ., Tokyo, Japan), and Makio Kashino (NTT Commun. Sci. Labs., Nippon Telegraph and Telephone Corp., Kanagawa, Japan)

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by an impaired ability to communicate and restricted and repetitive patterns of behavior. Sensory abnormality, a trait included in restricted patterns of behavior, is often observed in individuals with ASD. This study used an auditory illusion called "verbal transformation (VT)" to examine abnormal auditory/speech processing in ASD. In VT, a series of subjective perceptual changes occurs when subjects listen to a repeating word without a pause. The characteristic of VT is that, physically, the stimulus is the same repeated word but the perceptual change differs for each person. In this point, VT could be useful for demonstrating certain ASD traits such as favoring repetition. Twenty-four ASD adults and 24 neurotypical controls listened to three samples of repeated Japanese words for five minutes, and reported their percepts verbally whenever they experienced perceptual changes. As for the number of perceptual changes and that of reported forms, we found no significant difference between the ASD group and controls. However, the phonetic variations of form changes were different, and ASD group showed more phonetic variation than controls. We interpret the results from the viewpoint of unusual adaptation in auditory and/or phonetic processing.

1:45

5pSCa4. Listening effort and accent in competing-talker speech recognition by native and non-native listeners: Electrophysiological measures of auditory and lexical processing. Paul Iverson, Emma Brint, Jieun Song, and Mengfan Wu (Univ. College London, 2 Wakefield St, London WC1N 1PF, United Kingdom, p.iverson@ucl.ac.uk)

Listening effort modulates auditory processing for speech, such that attention to a target talker in the presence of a distractor increases the entrainment of cortical activity to the target speech amplitude envelope. Similarly, lexical-processing effort can be measured in EEG recordings, with more extensive lexical search being measured as a more negative N400. The present study compared effort at both levels for native speakers of southern British and Glaswegian English, and Chinese second-language English speakers. Listeners heard sentences spoken in southern British and Glaswegian accents, either with single or competing talkers. The results demonstrate that competing talkers and difficult accents have similar effects at a lexical level, increasing the amount of lexical search for native listeners and disrupting lexical search for non-native listeners. In contrast, cortical-acoustic entrainment was higher for targets than distractors in competing-talker conditions, but was not modulated by accent; entrainment was also similar for different language groups in the present study, although language-background differences in entrainment can be found in easier conditions with spatial cues to segregation. Accent may thus have little role in the

auditory segregation of competing talkers, despite the fact that accent and listener background affects intelligibility and lexical access.

2:00

5pSCa5. Adaptive speech modifications and its effect on communication effectiveness in complex acoustic environments. Nandini Iyer, Eric Thompson, Kelly Stillwagon, Zachariah Ennis, Abbey Willis, and Brian Simpson (Air Force Res. Lab., 2610 Seventh St., Bldg 441, Wpafb, OH 45433, Nandini.Iyer.2@us.af.mil)

Previous studies of communication in noisy environment have shown that interlocutors, when residing in the same acoustic environment, continuously adapt their speech patterns in order to overcome communication difficulties (Cooke et al., 2014; Lindblom, 1990). However, little is known about the impact on communication when individuals are in disparate acoustic environments and if/how these individuals would adapt their speech to support effective communication. In this study, eight pairs of interlocutors were required to completed the "spot-the-difference" Diapix task (Van Engel *et al.*, 2010) while being exposed to one of 2 acoustic environments presented at 85 dB SPL over headphones: 1) sparse (2-talker) babble; or 2) dense (8-talker) babble. In some conditions, both individuals were in similar acoustic environments while in others they were in different acoustic environments. A control condition, where both interlocutors communicated in quiet environments was also included. Preliminary results showed that the time for task completion increased when listeners communicated in a dense rather than sparse babble, thus indicating decreased communication efficiency. Acoustic characteristics of the speech produced in these adverse environments suggests that talkers modulate their speech for their own listening environment and not necessarily to meet the needs of their communication partners.

2:15

5pSCa6. The contribution of acoustic complexity to speech-specific processing. Faith Chiu and Jyrki Tuomainen (Speech, Hearing and Phonetic Sci., Univ. College London, Rm. 326, Chandler House, 2 Wakefield St., London WC1N 1PF, United Kingdom, Faith.Chiu.11@ucl.ac.uk)

Speech segments are delivered in an acoustically complex signal. This paper examines the role of acoustic complexity in speech-specific processing in the brain. Mismatch Negativity (MMN) is a deviance detection event-related response elicited by both speech and non-speech stimuli. Speech-specific MMN can further be sensitive to segment-driven categorisation processes (Näätänen *et al.*, 1997). Existing research on the neural basis of speech processing often employs only pure tones as a non-speech control (Schofield et al, 2009). We investigate the specific contribution of acoustic complexity to speech-specific processing using a passive oddball task to elicit MMN in three conditions: speech, spectrally rotated speech, and pure tones. EEG event-related potentials were examined for scalp-based observations and source connectivity; source activity at the level of pyramidal cells in neuronal groups was modeled as dynamic states (Kiebel *et al.*, 2007). Results indicate left lateralization only for acoustically complex sounds; both speech and rotated speech show within-region cortical coupling in left Heschl's gyrus. However, an increase in feedforward activity towards left superior temporal gyrus is only observed in speech but not rotated speech. Pure tone stimuli do not elicit left lateralization. We conclude that acoustic complexity is a necessary but insufficient condition for speech-specific processing.

2:30

5pSCa7. A study on the effect by adding reverberation to speech-like masking sound. Jen W. Tang, Jacky Wan, and Yusuke Hioka (Dept. of Mech. Eng., Univ. of Auckland, Privatebag 92019, Auckland 1142, New Zealand, jtan160@aucklanduni.ac.nz)

Problems arising from the acoustical privacy point of view in public spaces have been known to be an issue. The lack of acoustical privacy has been known to affect the human's health both physically and psychologically, thus keeping the acoustical privacy in public spaces will significantly reduce social loss. Masking is the most commonly practically used technique to make a target speech unintelligible to the unintended listeners

without needing to install any physical structures. Time-reversed speech has been known to effectively mask information for speech privacy applications; however, the annoyance and distraction caused by the time-reversed speech is known to be higher than other masking sound. This study explores a solution to compromise the suggested problem by adding a reverberant effect to a time-reversed speech. Subjective listening tests have been conducted to measure the intelligibility of target speech, annoyance and distraction caused by the masking sound. The experimental results suggest that adding artificial reverberation to a time-reversed speech has a significant effect to reduce the annoyance level while maintaining the masking effectiveness of the original time-reversed speech. A trend was also observed that the addition of artificial reverberation could reduce the level of distraction caused by the masking sound.

2:45

5pSCa8. Estimating realistic speech levels for virtual acoustic environments. Jorg M. Buchholz (Australian Hearing, National Acoust. Labs., 126 Greville St., Chatswood, NSW 2067, Australia, jorg.buchholz@nal.gov.au), Adam Weisser (Linguist, Audiol., Macquarie Univ., NSW, Australia), Tim Beechey, James Galloway, Gitte Keidser, and Katrina Freeston (Australian Hearing, National Acoust. Labs., Macquarie University, NSW, Australia)

Speech intelligibility is commonly assessed in rather unrealistic acoustic environments at negative signal-to-noise ratios (SNRs). As a consequence, the results seem unlikely to reflect the subjects' experience in the real world. To improve the ecological validity of speech tests, different sound reproduction techniques have been used by researchers to recreate field-recorded acoustic environments in the laboratory. Whereas the real-world sound pressure levels of these environments are usually known, this is not necessarily the case for the level of the target speech (and therefore the SNR). In this study, a two-talker conversation task is used to derive realistic target speech levels for given virtual acoustic environments. The talkers communicate with each other while listening to binaural recordings of the environments using highly open headphones. During the conversation their speech is recorded using close-talk microphones. Conversations between ten pairs of young normal-hearing talkers were recorded in this way in 12 different environments and the corresponding speech levels were derived. In this presentation, the methods are introduced and the derived speech levels are compared to results from the literature as well as from real sound-field recordings. The possibility of using this technique to generate environment-specific speech material with realistic vocal effort is discussed.

3:00–3:20 Break

3:20

5pSCa9. Effect of audibility, frequency region, and hearing loss on better-ear glimpsing. Baljeet Rana and Jörg M. Buchholz (National Acoust. Labs. (NAL), Level 5, Australian Hearing Hub 16 Univ. Ave., Sydney, NSW 2109, Australia, baljeet.rana@nal.gov.au)

Better-ear glimpsing (BEG), which utilises short-term interaural level difference cues (ILDs), helps understanding speech in noise. Since ILDs are mainly available at high frequencies where hearing loss is usually strongest, hearing-impaired (HI) listeners cannot take full advantage of BEG. One possible solution is to provide artificially generated ILDs at low and mid frequencies where hearing loss is usually less pronounced. Rana and Buchholz [JASA, 2016] showed that HI listeners can take advantage of such artificial ILDs, but their received benefit in speech intelligibility was still smaller than for normal-hearing (NH) listeners. To understand how far this difference can be explained by differences in audibility, speech recognition thresholds (SRTs) were measured in four frequency regions (low, mid, high, and broadband) at four audibility levels (0, 10, 20, and 30 dBSL) in 10 NH and 10 HI listeners. SRTs were measured for BKB like sentences in the presence of vocoded speech distractors. The difference between co-located and spatially separated SRTs was used as a measure of BEG. Results revealed that providing equal audibility removed the performance differences between groups. Moreover, the effect of BEG was roughly frequency independent and increased by about 2 dB for each 10 dB increase in sensation level.

5pSCa10. The role of modulation characteristics on the interaction between hearing aid compression and signal-to-noise ratio. Paul Reinhart (Northwestern Univ., 2240 Campus Dr., Evanston, IL 60208, preinhart@u.northwestern.edu), Pavel Zahorik (Univ. of Louisville, Louisville, KY), and Pamela Souza (Northwestern Univ., Evanston, IL)

Hearing aids use compression amplification to provide greater gain to lower amplitude sounds than higher amplitude sounds. When there are two signals (i.e., speech-in-noise) hearing aids alter the signal-to-noise ratio (SNR) by disproportionately increasing the lower amplitude signal compared to the higher amplitude signal. The purpose of the present study was to examine the relationship between hearing aid compression and SNR alteration along a continuum of signal modulation characteristics. Sentence stimuli were combined with either 1-, 2-, or 6-talker modulated noise at a range of SNRs. Speech-in-noise signals were processed at a range of reverberation times with the speech and noise signals co-located. Stimuli were then processed at a range of compression release times (RT) using a hearing aid simulation which mimicked six-channel hearing aid processing. Signals were acoustically analyzed using a digital inversion method which employed phase cancellation to separate speech and noise signals at the output of the compressor. This enabled us to quantify changes to SNR as a result of compression. Results indicated that shorter RTs altered SNR more than longer RTs; however, reverberation reduced or eliminated the differences among RTs. Additionally, less modulated maskers lead to poorer output SNRs than modulated maskers. [Work supported by NIH.]

5pSCa11. Real-time simulation of hearing impairment: Application to speech-in noise intelligibility. Nicolas Grimault (Univ. Lyon 1, CNRS UMR 5292 CRNL, UMR CNRS 5292 CRNL, Univ. Lyon 1 50 av T Garnier, Lyon cedex 07 69366, France, nicolas.grimault@cnrs.fr), Etienne Parizet (INSA-Lyon, Laboratoire Vibrations Acoustique, Villeurbanne, France), Alexandra Corneyllie (Univ. Lyon 1, CNRS UMR 5292 CRNL, Lyon, France), Laurent Brocolini (INSA-Lyon, Laboratoire Vibrations Acoustique, Villeurbanne, France), Richard Weyn, and Samuel Garcia (Univ. Lyon 1, CNRS UMR 5292 CRNL, Lyon, France)

Accurately simulating cochlear hearing loss in real time is of paramount importance for various applications, including easier testing of the impact of cochlear damage on auditory perception in real-life situations, evaluating sound qualities as perceived by hearing-impaired listeners, and pedagogical demonstrations. Following an idea originally proposed by Irino and colleagues, a real-time hearing-loss simulator was developed. Using an inverse compressive gammarchirp filters analysis, the simulator aims to replicate the increase of hearing thresholds, the level-dependent widening of auditory filters, and modification of the loudness curve (recruitment). The characteristics of the simulator, and basic psychoacoustical measurements performed with and without the simulator (absolute thresholds, auditory filters estimated with notch-noise data, cochlear input/output functions estimated with temporal masking curves) will be presented. Finally, an experiment comparing the intelligibility of speech in noise for real and simulated hearing-impaired listeners with a French version of the Four Alternative Auditory Feature Test (FAAF) will be presented. On average, the results of normal-hearing listeners using the simulator were found to be similar to the results obtained in individuals with hearing loss of presumed cochlear origin, suggesting that the simulator is valid.

5pSCa12. Within-listener comparisons of vowel confusions in pediatric, bilateral cochlear implantees. Julie A. Bierer (Speech and Hearing Sci., Univ. of Washington, 1417 NE 42nd St., Box 354875, Seattle, WA 98105, jbierer@uw.edu) and Mishaela DiNino (Neurosci., Univ. of Washing, Seattle, WA)

Tremendous variability in speech perception abilities exists in children and adults who wear cochlear implants (CIs). Preliminary data in adults suggests that variability in the interface between CI channels and auditory neurons can influence spectral resolution and thus vowel perception. One method for assessing the electrode-neuron interface is by using focused electrical fields to measure thresholds. A quick method for measuring

thresholds across the array has recently been validated in adults. Here, the findings are extended to children who were bilaterally implanted sequentially. These children have different etiologies than the typically studied adults and different durations of auditory deprivation and implant use between their two ears. By comparing within listeners, the role of cognitive or linguistic factors is minimized. Within ears, the confusions listeners made were consistent between conditions of quiet and 4-talker babble background noise at a +10 dB signal-to-noise ratio. For individual children, however, confusion patterns differed between the ears, suggesting that time with the implant and the electrode-neuron interface may indeed play a role in vowel confusions, as opposed to more global perception factors. Finally, the pattern of high-threshold regions corresponds to predictable vowel confusions, similar to those observed in adults.

5pSCa13. Cochlear implant listener vowel identification performance and confusion patterns with selective channel activation programs. Mishaela DiNino, Matthew B. Winn, and Julie A. Bierer (Dept. of Speech and Hearing Sci., Univ. of Washington, 1417 NE 42nd St., Seattle, WA 98105, MDiNino@u.washington.edu)

Cochlear implants (CIs) restore auditory perception to profoundly deaf individuals, yet identification of speech sounds remains limited by degraded or warped spectral representations. The present study aimed to measure the ways in which the perception of vowels (distinguished primarily by their spectra) is affected by selective activation of high or low quality electrode channels. Individualized experimental processing programs were created for eight CI users by selectively deactivating channels with poor electrode-neuron interfaces (identified by high auditory perception thresholds with focused stimulation), and reallocating frequencies to remaining electrodes ("High Off"). A contrary program in which channels with better electrode-neuron interfaces were deactivated ("Low Off") was created for each participant for comparison. A program with all channels activated ("All") served as a control. CI users performed a vowel recognition task with each experimental program. Overall percent correct did not change significantly across programs. However, perceptual distance and perceptual vowel space analyses indicated large differences in vowel confusion patterns between listening with "All," "High Off," and "Low Off" programs within individual CI users. These results suggest that vowel perception is dramatically altered by CI channel deactivation and frequency reallocation, which is not evident solely from average identification performance.

5pSCa14. Perception of low-pass filtered speech in hearing-impaired children, with and without cochlear dead regions and children with normal hearing. Alicja N. Malicka (School of Health and Rehabilitation Sci., The Univ. of Queensland, St. Lucia, QLD 4067, Australia, a.malicka@uq.edu.au), Kevin J. Munro (Manchester Ctr. for Audiol. and Deafness, The Univ. of Manchester, Manchester, United Kingdom), Thomas Baer (Dept. of Experimental Psych., Univ. of Cambridge, Cambridge, United Kingdom), Richard J. Baker (Manchester Ctr. for Audiol. and Deafness, The Univ. of Manchester, Manchester, United Kingdom), and Brian C. Moore (Dept. of Experimental Psych., Univ. of Cambridge, Cambridge, United Kingdom)

There are reports that adults with high-frequency cochlear dead regions (CDRs) exhibit an enhanced ability to use audible low-frequency acoustic information. This results in better performance of participants with CDRs compared to those without CDRs on tasks where the speech is low-pass filtered with a cutoff frequency near the edge frequency of the CDR (Moore and Vinay 2009). This enhanced ability to use low-frequency information may be related to cortical plasticity induced by the presence of a CDR and may be stronger in children due to maximal plasticity of the central auditory pathways. The aim of this study was to determine if children (aged 7-15 years) with a high-frequency congenital hearing loss and CDRs also show enhanced ability to use low frequency information. Vowel-consonant-vowel nonsense speech stimuli were low-pass filtered at various frequencies, amplified to correct for any hearing loss and presented via headphones. The percentage of correctly identified consonants for each low-pass filtering condition were measured in each ear separately. The performance of children with CDRs was compared against performance of age matched children

without CDR, children with normal hearing and against the Speech Intelligibility Index prediction. Preliminary data do not support the hypothesis that children with high-frequency CDRs use low-frequency information more effectively than their no-CDR counterparts or normal hearing children.

4:50

5pSCa15. For some bilateral cochlear implantees, a second implant can cause interference rather than improved speech understanding in the presence of competing speech. Matthew Goupell, Olga Stakhovskaya (Hearing and Speech Sci., Univ. of Maryland, College Park, 0119E Lefrak Hall, College Park, MD 20742, goupell@umd.edu), and Joshua G. Bernstein (Walter Reed National Military Medical Ctr., Bethesda, MD)

The goal of bilateral cochlear implantation is to provide binaural-processing benefits, such as improved sound localization and speech understanding in background noise. A recent study [Bernstein *et al.*, *Ear Hear.* 37, 282-288] showed that bilateral cochlear-implant (BICI) listeners could receive up to 5 dB of binaural speech-understanding benefit when listening with two ears compared to one. However, one outlier demonstrated interference rather than improvement from the second ear, which warranted further investigation. We recruited nine BICI listeners with pre- or peri-lingual onset of deafness and/or a long duration of deafness in at least one ear. We hypothesized these listeners might also demonstrate interference in binaural speech understanding. They listened to a male target and a male interferer, both taken from the Coordinate Response Measure corpus, in monaural (target and interferer in one ear) and binaural (monaural target, diotic interferer) listening configurations. On average, listeners received 9 dB of interference from their second ear. For listeners with a much longer duration of deafness in one ear, interference patterns were asymmetrical (little interference in the

first implanted ear, extreme interference in the second). These data demonstrate that certain spatial listening configurations may hinder the advantages of bilateral implants for some users.

5:05

5pSCa16. Can spectral centroid explain voice pitch and vocal-tract length perception in normal-hearing and cochlear implant listeners? Etienne Gaudrain (Auditory Cognition and PsychoAcoust., CNRS UMR 5292, Lyon Neurosci. Res. Ctr., Université Lyon 1, UMCG, KNO, Huispostcode BB20, PO Box 30.001, Groningen 9700 RB, Netherlands, etienne.gaudrain@cnrs.fr)

Voice pitch (F0) and vocal-tract length (VTL) are two principal voice characteristics that play a major role in segregating voices in cocktail party situations. Cochlear implant (CI) listeners struggle with taking advantage of voice differences in such situations. Recent studies show that they have difficulties discriminating voices on the basis of these two cues. Yet, the mechanisms underlying perception of these cues in CI users, but also in normal hearing (NH) listeners, remain largely unknown. In CIs, F0 can be coded temporally and spectrally in the lower frequency channels, but some recent studies have suggested that spectral centroid (SC) could be used instead. VTL could be perceived through its effect on individual formants, but is also often likened to timbre, and as such, it has been argued that VTL perception might rely on SC. However, these assumptions tend to overlook the SC variability occurring in natural speech. In this study, it is shown how this variability would influence F0 and VTL JNDs in NH and CI listeners if they were based on SC, using a signal detection theory model. The results suggest that SC is unlikely to be a reliable cue for vocal F0 and VTL perception in natural speech.

FRIDAY AFTERNOON, 2 DECEMBER 2016

CORAL 3, 1:00 P.M. TO 5:30 P.M.

Session 5pSCb

Speech Communication: Speech Perception and Production by Clinical, Aging, and/or Developing Populations (Poster Session)

Jessica Alexander, Chair
Psychology, Centenary College of Louisiana, 2911 Centenary Blvd, Shreveport, LA 71104

All posters will be on display from 1:00 p.m. to 5:30 p.m. To allow contributors in this session to see the other posters, authors of odd-numbered papers will be at their posters from 1:00 p.m. to 3:15 p.m. and authors of even-numbered papers will be at their posters from 3:15 p.m. to 5:30 p.m.

Contributed Papers

5pSCb1. Perception of geminate consonants and devoiced vowels in Japanese by elderly and young listeners. Eri Iwagami, Takayuki Arai (Sophia Univ., 7-1, Kioi-cho, Chiyoda-ku, Tokyo 102-8554, Japan, alice.erieri@gmail.com), Keiichi Yasu (Res. Inst. of National Rehabilitation Ctr. for People with Disabilities, Tokyo, Japan), Emi Yanagisawa (Meiji Univ., Tokyo, Japan), and Kei Kobayashi (The Univ. of Auckland, Auckland, New Zealand)

The goal of this study is to compare the perception of geminate consonants and devoiced vowels by elderly and young Japanese speakers. This study consisted of two perceptual experiments (Exp. 1 and Exp. 2). In Exp.1, materials were nonsense words "ata" and "atta" having the structures: 'V_1CV_2', '$V_1C:V_2$', and the following three parameters were changed: 1) formant transition of V_1 (FT), 2) V_1 duration, and 3) closure duration. In Exp. 2, materials were 27 Japanese words whose vowels can be devoiced. The results of Exp. 1 showed that compared with young listeners the

perceptual boundary of geminate for elderly listeners shifted towards longer closure duration in the absence of FT. Moreover, the results of Exp. 2 showed that the misperception rate of young listeners was low with and without devoicing, whereas the rate was high with devoicing and low without devoicing in elderly listeners.

5pSCb2. Cognitive predictors of older adults' perception of difficult speech. Erin Ingvalson, Kaitlin Lansford, Valeriya Federova, and Gabriel Fernandez (School of Commun. Sci. and Disord., Florida State Univ., 201 W Bloxham St., Tallahassee, FL 32306, eingvalson@fsu.edu)

Recent work suggests that receptive vocabulary, over working memory capacity, may predict older adults' ability to perceive disordered speech, particularly when hearing acuity is poor (McAuliffe *et al.*, 2013). Further work in younger adults suggests that both receptive vocabulary and

3439   J. Acoust. Soc. Am., Vol. 140, No. 4, Pt. 2, October 2016

5th Joint Meeting ASA/ASJ   **3439**

cognitive flexibility may predict listeners' ability to perceive accented speech (Baese-Berk, *et al.*, 2015). Fifteen older adults transcribed phrases produced by five non-native English speakers and four talkers with dysarthria and completed measures on receptive vocabulary, auditory working memory, cognitive flexibility, and inhibitory control. Listeners were more accurate identifying speech produced by talkers with dysarthria than by talkers with foreign accents ($t(14) = 11.08$, $p < .001$). Mixed effect models indicated that hearing acuity, auditory working memory, and inhibitory control were significant predictors of both difficult speech conditions; cognitive flexibility was an additional predictor of older adults' ability to perceive dysarthric speech. In contrast to earlier work, receptive vocabulary was not a significant predictor of listeners' accuracy in either condition, nor were listeners' perception of the accented speech and the dysarthric speech correlated. These data add to the growing picture that difficult speech perception is a cognitively demanding task for older adults, regardless of hearing ability.

5pSCb3. Articulation rates of people who do and do not stutter during oral reading and speech shadowing. Rongna A (Res. Inst., National Rehabilitation Ctr. for Persons with Disabilities, 4-1 Namiki, Tokorozawa, Saitama 3598555, Japan, hohchahar-aruna@rehab.go.jp), Keiko Ochi (National Inst. of Informatics, Tokyo, Japan), Keiichi Yasu, Naomi Sakai, and Koichi Mori (Res. Inst., National Rehabilitation Ctr. for Persons with Disabilities, Tokorozawa, Saitama, Japan)

Purpose: Previous studies indicate that people who stutter (PWS) speak more slowly than people who do not stutter (PWNS), even in the fluent utterances. The present study compared the articulation rates of PWS and PWNS in two different conditions: oral reading and speech shadowing in order to elucidate the factor that affect the speech rate in PWS. *Method*: All participants were instructed to read aloud a text and to shadow a model speech without seeing its transcript. The articulation rate (mora per second) was analyzed with an open-source speech recognition engine "Julius" version 4.3.1 (https://github.com/julius-speech/julius). The pauses and disfluencies were excluded from the calculation of the articulation rate in the present study. *Results*: The mean articulation rate of PWS was significantly lower than that of PWNS only in oral reading, but not in speech shadowing. PWS showed a significantly faster articulation rate, comparable to that of the model speech, in shadowing than in oral reading, while PWNS did not change the articulation rate depending on the task. *Conclusion*: The group and task comparisons indicate that the ability of PWS to rapidly articulate is not impaired but somehow impeded during oral reading.

5pSCb4. Acoustics of palilalia: A case study. Christina C. Akbari (Commun. Disord., Arkansas State Univ., 4305 Bekah Dr. #8, Jonesboro, AR 72404, cakbari@astate.edu) and Amy Shollenbarger (Commun. Disord., Arkansas State Univ., State University, AR)

Palilalia is a rather rare speech disorder associated with a variety of etiologies including Parkinson's disease, stroke, epilepsy, and others. According to the literature, palilalia is characterized by repetitions of words, phrases, and/or sentences produced with an increasing rate and decreasing intensity. Because of the rare nature of this disorder, literature involving the characteristics is limited and somewhat contradictive. Although the original definition involves increasing rate and decreasing intensity, some researchers (i.e., Kent & Lapointe, 1982) have found different characteristics such as decreasing rate and increasing intensity when examining individual cases. This study adds to the literature regarding characteristics of palilalic speech in terms of change in rate and fundamental frequency from the original utterance to the first repeated train. This study utilized a single-subject design involving a 42 year old male who suffered an anoxic brain injury. Spontaneous speech samples were collected and analyzed. The results showed significant differences in terms of rate and fundamental frequency with the first repeated train in comparison to the original utterance. The rate was significantly faster and the fundamental frequency was significantly higher in the repeated utterance.

5pSCb5. Influence of clear speech on voice onset time and intelligibility of word-initial stop consonants in electrolaryngeal speech. Steven R. Cox, Lawrence Raphael (Commun. Sci. and Disord., Adelphi Univ., Hy Weinberg Ctr. Rm. 136, Adelphi University, Garden City, NY 11530, scox@adelphi.edu), and Philip C. Doyle (Health and Rehabilitation Sci., Western Univ., London, ON, Canada)

This study addressed the influence of clear speech (CS) on voice onset time (VOT) and intelligibility of word-initial (WI) stops in electrolaryngeal (EL) speech. Eighteen consonant-vowel-consonant words containing /p/, /t/, /k/, /b/, /d/, and /g/ in WI position were spoken by 10 laryngectomees in both habitual speech (HS) and clear speech (CS) conditions. Twelve, naive listeners transcribed a total of 4,320 words across HS and CS conditions of which 720 words (containing WI stops) were analyzed. Results indicate that when using CS, EL speakers produced stops with ~4 milliseconds of greater lag time compared to HS. Further, the intelligibility of WI stops improved 3.8% during CS. Finally, listeners omitted 2.22% more stops in WI position when using HS compared to CS. However, repeated measures analysis of variance revealed no significant effect of speaking condition on VOT or the intelligibility of stops. Collectively, these findings provide initial evidence that volitional attempts to make EL speech clearer via CS does not necessarily facilitate improved VOT or intelligibility for EL speakers.

5pSCb6. Patterns in speech errors among children with auditory processing disorder. Brooke DeVore (Dept. of Linguist and Cognit. Sci., Univ. of Delaware, 210 South College Ave., Newark, DE 19716, bdevore@udel. edu), Kyoko Nagao, Olivia Pereira (Ctr. for Pediatric Auditory and Speech Sci., Nemours/Alfred I. duPont Hospital for Children, Wilmington, DE), Julianne Nemith, Rachele Sklar (Dept. of Linguist and Cognit. Sci., Univ. of Delaware, Newark, DE), Emily Deeves (Dept. of Speech, Lang., and Hearing Sci., Univ. of Colorado Boulder, Boulder, CO), Emily Kish (Dept. of Commun. Sci. and Disord., West Chester Univ., West Chester, PA), Kelsey Welsh (Dept. of Commun. Sci. and Disord., Montclair State Univ., Montclair, NJ), and Thierry Morlet (Ctr. for Pediatric Auditory and Speech Sci., Nemours/Alfred I. duPont Hospital for Children, Wilmington, DE)

Auditory Processing Disorder (APD), a neurological hearing impairment, often leads to delays in speech development when diagnosed in childhood. It is unclear whether children with APD make articulation errors due to auditory processing deficits. The purpose of this study was to analyze the articulation error patterns of children diagnosed with APD. We examined speech samples from 26 children diagnosed with APD who participated in a previous research study involving speech perception testing. During this task, each child repeated each of 50 monosyllabic words presented to them in quiet. For each incorrect response, we examined the errors by type of articulation error (substitution, deletion, insertion, or distortion) and the syllable position (onset or coda). Speech perception scores were normal in quiet. However, 22 subjects responded incorrectly for one or more words, with substitution errors (40-50%) being the most prevalent type of errors. Articulation errors in the coda position (81%) were more common than errors in the onset position (19%). These results imply that children with APD, though generally performing well on speech tests in quiet, tend to display articulation issues not typically seen in normally developing children in the same age group. [Work supported by NIH COBRE Grants 2P20RR020173-06A1 and P30GM114736.]

5pSCb7. Exploiting the Ganong effect to probe for phonetic uncertainty resulting from hearing loss. Steven P. Gianakas and Matthew Winn (Speech & Hearing Sci., Univ. of Washington, 4217 NE 42nd St., Seattle, WA 98105, spgia5@uw.edu)

The "Ganong effect" is the tendency to perceive an ambiguous speech sound as a phoneme that would complete a real word, rather than completing a nonsense/fake word. For example, a sound that could be heard as either /g/ or /k/ is perceived as /g/ when followed by "ift" but perceived as /k/ when followed by "iss." Because the target speech sound (/g/ or /k/) is the

same across both environments, this effect demonstrates the influence of top-down processing rather than simply bottom-up processing in speech perception. We hypothesized that degradations in the auditory system (including simulations of hearing loss or cochlear implantation), speech stimuli will be rendered more ambiguous, and an increased Ganong effect should be observed. Participants heard three speech continua that varied by spectral cues of varying speeds, including fast (stop formant transitions), medium (fricative spectra), and slow (vowel formants). Stimuli were presented with clear spectral quality, or with varying amounts of spectral degradation using a noise vocoder. Responses were analyzed within stimulus pairs using binomial logistic regression, revealing increased Ganong effect with degraded speech. This test is proposed as a simple measure of how phonetic uncertainty can lead to an individual's increased reliance on top-down processing.

5pSCb8. Target-word identification in noise with formant enhancement for hearing-impaired listeners. Jingjing Guan and Chang Liu (Univ. of Texas at Austin, 1 University Station A1100, Austin, TX 78712, jane.guan@utexas.edu)

Degraded speech intelligibility in background noise is a common complaint of most listeners with hearing loss. The purpose of the current study is to investigate the effect of spectral enhancement of F2 on target-word identification in noise for the old listeners with hearing loss (HL) and with normal hearing (NH). Target words (e.g., name, color, and digit) were selected and presented based on the paradigm of coordinate response measure (CRM) corpus of English and Chinese versions. Identification tasks with original and F2-enhanced speech in the 2-talker and 6-talker babbles were designed for the two groups (HL and NH). Thresholds of word identification was measured by using adaptive up-down methods. Results showed that listeners with NH had better performance on word identification in noise than listeners with HL in almost all listening conditions. More importantly, thresholds of both groups were improved for enhanced speech signals. Compared with NH group, listeners with HL gained significantly greater benefits in the most challenging condition. These preliminary results suggested that speech sounds with F2 enhancement might improve listeners' ordinary speech perception in noise.

5pSCb9. Ideal binary masking based noise reduction and interleaved processing for bilateral cochlear implant users. Shaikat Hossain (Univ. of Texas at Dallas, 7220 Mccallum Blvd. #307, #307, Dallas, TX 75252, shossa3@gmail.com), Vahid Montazeri, and Peter F. Assmann (Univ. of Texas at Dallas, Plano, TX)

The present study investigates the interaction between ideal binary masking (IdBM) noise reduction and interleaved processing for bilateral cochlear implant (CI) users and normal-hearing (NH) listeners attending to vocoder simulations. IdBM decomposes a signal into time-frequency (T-F) bins and retains only regions where the target is higher than a local threshold (LC). IdBM benefits in CI users have been found to be limited due to factors such as current spread from neighboring electrodes. Interleaving channels across ears in bilateral CI users has been one approach to mitigating the effects of current spread. In the present experiments we tested NH listeners attending to vocoder simulations and CI users with IEEE sentences presented from different azimuths in speech-shaped noise at 5 dB signal-to-noise ratio and IdBM processing with LC values of 5 and -10 dB. Speech intelligibility was poorer for the interleaved condition in the presence of noise. As predicted, however, greater IdBM benefits were found for interleaved compared to non-interleaved processing for both NH and CI listeners. For CI users, the LC value of 5 dB led to greater improvements whereas the NH listeners had comparable benefits for both LC values. Findings have relevance for the clinical application of interleaved processing.

5pSCb10. Effect of dimension-switch ability in children with high-functioning Autism's perception of emotional prosody. Chieh Kao and Feng-Ming Tsao (National Taiwan Univ., No. 1, Sec. 4, Roosevelt Rd., Taipei 106, Taiwan, chiehkao.ckao@gmail.com)

Children with High-Functioning Autism's (HFA) sensitivity to emotional prosody in speech sounds, was long under debate. The stiffness

prohibiting them to switch between diverse features might be one of the factors that make them difficult to understand emotional prosody in daily communications. Present study explored how dimension-switch ability in identifying lexical tone would modulate children's perception of emotional prosody, since both prosody and lexical tone are conveyed by pitch height and contour. Mandarin-speaking HFA (n = 22) and typically developing (TD, n = 16) children aged 7 to 12 first classified emotions of 24 emotion-prosody congruent/incongruent words. Second, they identified lexical tone in a categorical perception task, and then received a dimension switch task, which asked them to choose the wrong answer. Two groups of children performed similarly in these tasks. HFA children identified basic emotions well as TD children even under the incongruent word pairs, which could be explained by their even performances in dimension switch task. Despite the similarity, there was a correlation between the accuracy rate of emotion classification and lexical tone perception in HFA children, which was not seen in TD children. This showed that HFA children used correlated mechanisms in understanding emotional prosody and lexical tone.

5pSCb11. Auditory frequency discrimination in preschool and school-age children with specific language impairment. Huei-Mei Liu (Special Education, National Taiwan Normal Univ., 162 Ho-Ping East Rd., Sec. 1, Taipei 106, Taiwan, liumei@ntnu.edu.tw) and Feng-Ming Tsao (Psych., National Taiwan Univ., Taipei, Taiwan)

The view of "perceptual processing deficits" in children with specific language impairment (SLI) emphasized that the detection and discrimination of basic acoustic components contributed to their difficulties in higher linguistic process. The purpose of this study was to examine the auditory frequency discrimination in Mandarin-speaking children with SLI across ages. Eighteen 5-year-old SLI and 18 age-matched controls, 20 8-year-old SLI and 20 controls participated. All the participants were assessed by the standardized intelligence and language tests. In the computerized auditory frequency discrimination task, children were asked to do AX discrimination of 1000-1030Hz and 1000-1070 Hz in conditions with and without noise masking. A mixed ANOVA was used to examine the effects of group, frequency difference, and masking conditions. The results showed that both preschool and school-age children with SLI performed significantly poorer on the frequency discrimination compared to the controls, especially in the masking condition. The preschoolers performed poorer on the frequency difference of 30 Hz than 70 Hz, but the school-age children performed similarly on both frequency differences. Individuals' frequency discrimination were positively correlated with their language performance across two ages. These results suggest that deficits in frequency discrimination in children with SLI are important to their language difficulties.

5pSCb12. Do clear-speech consonants always provide the best intelligibility for hearing-impaired listeners? Matthew J. Makashay (Army Hearing Program, Army Public Health Ctr. (Provisional), 5158 Blackhawk Rd., Aberdeen Proving Ground, MD 21010, matthew.j.makashay.civ@mail.mil), Nancy P. Solomon, and Van Summers (Audiol. and Speech Pathol. Ctr., Walter Reed National Military Medical Ctr., Bethesda, MD)

Hearing-aid amplification enhances speech intelligibility for many hearing-impaired (HI) listeners in quiet, but often provides no clear benefit in noise. Requesting that talkers use clear speech is one strategy to overcome these listening challenges. Paradoxically, one feature of clear speech is a shift to higher frequencies, which may move speech energy into a frequency range that is inaudible or has more distortion for certain HI listeners. Conversely, casual conversational speech may shift speech energy into a lower frequency range that is more audible or has less distortion. This study examined the intelligibility of 21 casually- and clearly-spoken American English coda consonants in nonsense syllables for 9 aided normal-hearing and 18 aided HI listeners. As expected, most clear-speech consonants yielded higher recognition scores. However, certain phonological processes common in casual speech, namely affrication and palatalization, generated significantly higher scores than their clear counterparts for some HI listeners in noise. These results have implications for coaching conversational partners of aided HI listeners. [The views expressed in this abstract are those of the authors and do not reflect the official policy or position of the Department of the Army, Department of Defense, or the U.S. Government.]

5pSCb13. Bite block effects on vowel acoustics in talkers with amyotrophic lateral Sclerosis and Parkinson's Disease. Antje Mefferd and Mary Bissmeyer (Dept. of Hearing and Speech Sci., Vanderbilt Univ. Medical Ctr., 8310 Medical Ctr. East, Nashville, TN 37232, antje.mefferd@vanderbilt.edu)

Amytrophic Lateral Sclerosis (ALS) affects predominantly the tongue and leaves the jaw relatively spared. By contrast, Parkinson's disease (PD) is thought to affect the motor system globally with minimal jaw movements potentially further limiting tongue articulatory movements. However, currently only few studies have directly investigated articulator-specific contributions to deviant speech acoustics in talkers with dysarthria. In this study, a 10 mm bite block was used to decouple tongue and jaw and delineate their specific contributions on vowel acoustics in talkers with distinctly different impairment profiles. Acoustic vowel contrast was examined during sentence repetitions produced by talkers with ALS, PD, and controls under jaw-free and jaw-fixed conditions. Preliminary findings of seven speakers per group revealed that acoustic vowel contrast was significantly lower during the jaw-fixed than the jaw-free condition in ALS. However, in PD acoustic vowel contrast tended to be greater when the jaw was fixed. For controls, no difference was found between the two conditions. Further, compared to controls acoustic vowel contrast was significantly reduced during the jaw-fixed condition in ALS and during the jaw-free condition in PD. Outcomes suggest that the jaw promotes acoustic vowel contrast in ALS but may hinder it in PD.

5pSCb14. Changes in lateralized /s/ production after treatment with opti-speech visual biofeedback. Rebecca L. Mental (Case Western Reserve Univ., 2744 Euclid Heights Blvd., Apt. 3, Cleveland Heights, OH 44106, rlm142@case.edu), Holle Carey (Vulintus, Dallas, TX), Nolan Schreiber, Andrew Barnes, Gregory S. Lee, and Jennell Vick (Case Western Reserve Univ., Cleveland, OH)

Opti-Speech is a visual biofeedback software for the treatment of speech sound placement errors. It utilizes Northern Digital Wave platform for electromagnetic articulography. Opti-Speech tracks the movement of five sensors on a talker's tongue to animate a 3D tongue avatar that moves in real time with the talker's own tongue. The avatar is viewable from multiple angles, and virtual targets can be created to guide the participant's speech movements. Opti-Speech has found initial success in small feasibility studies (Katz and Mehta, 2015; Vick et al., 2016). The two participants described in this study are part of a larger scale clinical trial to evaluate efficacy. Both adolescent males presented with the lateralized /s/ speech error. Their intensive biofeedback treatment schedule included two one-hour sessions per day over five sequential days. Data analyses include acoustic measures (spectral mean and kurtosis of the /s/) and trained perceptual judgments of speech sound accuracy on a 16 item probe list from baseline and across all treatment sessions. Overall impressions of the utility of this biofeedback system as a treatment tool will be discussed from the perspective of both the treating clinician and the participants.

5pSCb15. Phonological deficits in Mandarin-speaking congenital amusics. Yun Nan, Wei Tang, and Qi Dong (State Key Lab. of Cognit. Neurosci. and Learning, Beijing Normal Univ., 204 Office, Brain Imaging Ctr., 19 Xin-Wai St., Hai-Dian District, Beijing 100875, China, nany@bnu.edu.cn)

Congenital amusia is a neurodevelopmental disorder that mainly affects pitch processing in music and speech. Recent research reveals possible phonological deficits among amusic individuals from non-tonal language speakers. In a tone language such as Mandarin, only a subgroup of amusic individuals showed speech tone deficits. In the present study, we examined whether a similar phonological deficit exists in Mandarin amusics and whether the phonological deficits are confined to the subgroup of amusics who showed speech tone deficits. An simple identification task targeted at either vowel, tone, or a combination of vowel and tone for 20 monosyllabic

words was administered to three groups of Mandarin speakers: amusics with (tone agnosics) and without tone deficits (pure amusics), and matched controls. Tone agnosics lagged significantly behind both the pure amusics and controls when identifying tone and the combination of vowel and tone, but performed similarly to the latter two groups when identifying vowel. However, pure amusics performed comparable with the control group for identification of vowel, tone, and the combination of vowel and tone. These results suggest that only the tone agnosics were impaired in phonological processing and their phonological difficulties are mainly due to the speech tone deficits.

5pSCb16. Acoustic measures versus crowdsourced listeners' ratings: Evaluating the efficacy of treatment for hypokinetic dysarthria. Christopher Nightingale (Speech-Lang. Pathol., Hawaii State Dept. of Education, 2450 Prince Edward St., Apt. 204A, Honolulu, HI 96815, cn883@nyu.edu), Michelle T. Swartz (Communicative Sci. & Disord., New York Univ., Beesleys Point, NJ), Lorraine A. Ramig (Speech, Lang., and Hearing Sci., Univ. of Colorado-Boulder, Boulder, CO), and Tara McAllister Byun (Communicative Sci. & Disord., New York Univ., New York, NY)

Clinical interventions for speech disorders aim to produce changes that are not only acoustically measurable or perceptible to trained professionals, but are also apparent to naïve listeners. Due to the difficulty of collecting ratings from a suitably large sample, few researchers evaluate speech interventions relative to this criterion. Crowdsourcing technologies could enable a shift toward a more naturalistic standard to evaluate speech interventions. This project compared 35 naïve crowdsourced listeners' ratings against acoustic measures of speech samples collected from patients with hypokinetic dysarthria secondary to Parkinson's Disease. The data come from a published efficacy study (Sapir et al., 2007) that documented significant acoustic changes after a period of intensive treatment, Lee Silverman Voice Treatment. Specifically, the ratio of mean F2 in /i/ versus /u/, calculated for each subject, was greater in post-treatment relative to pre-treatment samples. Mixed-effects logistic regression indicated that words elicited post-treatment, presented in randomly ordered pairs with words elicited pre-treatment, were significantly more likely to be rated "more clear" (beta = 1.24, SE =.46, p <.01). This result supports the original conclusion of Sapir et al. (2007); it also supports the validity of crowdsourcing as a means to obtain ratings of disordered speech data.

5pSCb17. Prosodic changes in Parkinson's disease. Sona Patel (Speech-Lang. Pathol., Seton Hall Univ., 400 South Orange Ave., South Plainfield, NJ 07080, sona.patel@shu.edu), Sabiha Parveen (Commun. Sci. and Disord., Oklahoma State Univ., Stillwater, OK), and Supraja Anand (Commun. Sci. and Disord., West Chester Univ., West Chester, PA)

Parkinson's disease often results in a hypokinetic dysarthria, causing well-known effects on speech and voice. Speech from individuals with Parkinson's is typically studied through listening tasks (i.e., by family members and naïve listeners) and is known to affect intelligibility. However, the exact speech changes causing degradations in speech quality have not been defined. Previous investigations have focused on standard measures of mean fundamental frequency, intensity, and rate. We hypothesize that speech changes are present at the prosodic level in addition to at the acoustic level. To test this hypothesis, we collected spontaneous conversational speech samples from 15 individuals with Parkinson's and 13 age-and-gender matched neurologically healthy speakers (Control group). Several parameters were extracted from the speech signal to describe rate, phrasing, and stress in addition to fundamental frequency and intensity. These measures were compared with listener ratings of intelligibility performed by 12 naïve listeners. Listeners rated the speech samples using a visual analog scale. Preliminary findings suggest that prosody is impaired in Parkinson's disease. Statistical analysis of variance between the speaker groups (Parkinson's vs. Controls) are reported in addition to linear regressions with listener ratings. Our findings can facilitate in early diagnosis and the development of targeted speech therapy for individuals with Parkinson's.

5pSCb18. Effect of F0 on the intelligibility of emotional speech in noise for younger and older listeners. Jana Besser (Dept. of Otolaryngology-Head and Neck Surgery, Section Ear & Hearing and EMGO Inst. for Health and Care Res., VU Univ. Medical Ctr., Zurich, Switzerland), Margaret K. Pichora-Fuller (Psych., Univ. of Toronto, 3359 Mississauga Rd., Mississauga, ON L5L 1C6, Canada, k.pichora.fuller@utoronto.ca), S T. Goverts, Sophia E. Kramer, and Joost M. Festen (Dept. of Otolaryngology-Head and Neck Surgery, Section Ear & Hearing and EMGO Inst. for Health and Care Res., VU Univ. Medical Ctr., Amsterdam, Netherlands)

This study investigated if F0 variability could explain why intelligibility in noise is better for speech spoken to portray fear compared to emotionally neutral speech. Word recognition accuracy was measured for stimuli produced with neutral vocal emotion, for intact stimuli portraying fear, and for six versions of the fear stimuli with varying reductions in F0 variability. Younger and older adults were tested in speech-spectrum noise and two-talker babble. Younger adults outperformed older adults. As F0 variability in the fear stimuli was reduced, performance in speech-spectrum noise for both age groups decreased and approached performance for the neutral stimuli. In the two-talker babble, even when F0 variability was most reduced, performance remained higher for the fear stimuli than the neutral stimuli, especially for older adults. The mean F_0 for the fear stimuli was higher than for the neutral stimuli, while mean F0 of the neutral stimuli and the two-talker masker were similar. Thus, although the greater F0 variability in the fear stimuli confers an advantage to both age groups in speech-spectrum noise, F0 mean and variability both contribute to the effect of emotion on intelligibility in two-talker babble, especially for older listeners.

5pSCb19. Vocal fold vibration and voice quality in the elderly. Robin A. Samlan (Speech, Lang., & Hearing Sci., Univ. of Arizona, P.O. Box 210071, Tucson, AZ 85721, rsamlan@email.arizona.edu), Melda Kunduk (Commun. Sci. & Disord., Louisiana State Univ., Baton Rouge, LA), and Takeshi Ikuma (Otolaryngol., LSU Health Sci. Ctr., New Orleans, LA)

The vocal fold vibration patterns of incomplete glottal closure, aperiodicity, and asymmetry are characteristic of older adults with normal voice quality. Given that these patterns are also reported in older adults with voice concerns (i.e., age-related dysphonia, presbylarygis, or presbyphonia), the question of whether they relate to voice quality becomes relevant to providers making evaluation and management decisions. METHOD: Thirty-eight adults aged 70 years and older underwent laryngeal high-speed video examination at 4000 frames per second. Analysis included measurements of the glottal closure over time (i.e., open quotient, relative glottal gap), spectral noise (i.e., harmonics-to-deterministic-noise ratio, harmonics-to-noise ratio), and symmetry (i.e., amplitude symmetry, phase symmetry). Multiple regression will be used to compare the high-speed measures to cepstral peak prominence of the acoustic signal and perceived voice quality. Results will be interpreted in terms of how glottal closure and vibratory periodicity and symmetry affect voice quality in elderly speakers. Implications for clinical management will be discussed.

5pSCb20. Perceptual differentiation of African American and European American children: Accuracy, confidence, and reaction time. Emily Simmons, Dalila Salas, Nicole Marsh, Julia Licata, and Sonja Trent-Brown (Hope College, 35 E 12th St., Holland, MI 49423, emily.simmons@hope.edu)

Previous research has demonstrated that adult speakers can be perceptually differentiated with respect to gender and ethnicity (Thomas & Reaser, 2004). We examine how age, gender, and ethnicity of child speakers, ages 8-12, affect the perceptual accuracy of adult listeners, determining the point at which perceptual accuracy meets the level for adult speakers. Undergraduates listened to forward and reversed recordings of European American and African American children producing /h-vowel-d/ words (/i,ɪ,e,ɛ,æ,ʌ,ɑ,ɔ,o,ʊ,u,ɝ/) and sentences ("*I hear the sound of /h-V-d/ some more*") in General American English. Listeners identified speaker ethnicity and gender and rated their confidence. Reaction times were also measured for identification and confidence ratings. We expect greater accuracy, faster reaction times, and higher median ratings for forward blocks, sentences, and female speakers. We do not expect there to be a difference in identification of European American voices versus African American voices. We also

anticipate that accuracy of identification will improve as the age of the speaker increases. Implications of this work can be extended for individuals who work with children and in speech-related professions to reduce biases that occur as a result of personal language experience.

5pSCb21. The use of ultrasound to determine the most effective placement cues for /s/ during articulation therapy. Kathleen Siren (Speech-Language-Hearing Sci., Loyola Univ. Maryland, 4501 North Charles St., Baltimore, MD 21210, ksiren@loyola.edu)

Although two distinct normal tongue configurations for /s/ have been documented for many years (Borden & Gay, 1978), estimates of the frequency of tongue tip up vs. tongue tip down /s/ articulation in the general population vary. Therapy procedures often focus exclusively on one or the other of the two tongue postures. Given that /s/ is one of the most frequently misarticulated sounds in English, clinicians may need guidance in determining which tongue placement is most natural for each client. In recent years, portable ultrasound devices have become more widely accessible for use in speech therapy. In such cases, ultrasound has generally been used as a visual feedback technique for clients (Bernhardt, Gick, Bacsfalvi & Adler-Bock, 2005; Cleland, Scobbie & Wrench, 2015). The current study investigates the use of ultrasound imaging as a preliminary information gathering tool for clinicians to guide decisions regarding /s/ therapy techniques and procedures. In this investigation, students utilized ultrasound to attempt to determine tongue tip configuration during /s/ production for various individuals. Results indicate that ultrasound can provide clinicians with enough initial information to determine the most natural tongue configuration for /s/ production for most clients, thus potentially reducing therapy time and increasing therapy effectiveness.

5pSCb22. Effect of speech tasks on cepstral measures of articulation. Supraja Anand, Mark D. Skowronski (USF, 4202 East Fowler Ave., PCD 1017, Tampa, FL 33620, supraja1984@gmail.com), and Rahul Shrivastav (UGA, Athens, GA)

Hypokinesia in articulatory movements is one of the predominant characteristics in individuals with Parkinson's disease (PD). Past research quantified such hypokinesias through measures of formant frequencies (F1 and F2) and/or derivative measures such as vowel space area. Extractions of these acoustic measures are labor intensive, time consuming and require phonetic segmentation. In this study, we characterize acoustic measures of articulatory range, rate and acceleration using human factor cepstral coefficients (HFCCs) across speech tasks (e.g. syllable repetition, sentence reading and monologues). Cepstral measures of articulation can be readily derived from entire sentences in an automated, time/computationally efficient method. We hypothesize that these cepstral measures will show a significant effect of group and speech task. Twelve individuals with clinical diagnosis of idiopathic PD and age/gender matched healthy controls repeated syllables at a fast rate, read some sentences, and provided monologue samples. Articulatory range, rate and acceleration were studied using cepstral coefficient standard deviation sum (CC SDS), delta (Δ), and delta delta (ΔΔ) coefficients, respectively. Preliminary findings suggest that temporal derivatives differentiated the groups better in the syllable repetition task and CC SDS revealed trends in the reading and monologue tasks, showing promise for differential and early diagnosis.

5pSCb23. Mathematical error in the American Medical Association and American Academy of Otolaryngology-Head and Neck Surgery percent binaural hearing impairment calculation. Ashley N. Clark and Ron Leavitt (Audiol., Corvallis Hearing Ctr., 1025 NW Ninth St. Ste. D, Corvallis, OR 97330, researchchc@gmail.com)

Percent hearing loss calculation forms the basis of computation for those whose hearing has been damaged in industry or in the military. As such medical professionals are admonished to use the best science available when making these important calculations. Unfortunately, current computational methods do not meet current high standards of scientific evidence and are riddled with mathematical and psychoacoustic errors. A number of these errors will be discussed in this poster session.

3443   J. Acoust. Soc. Am., Vol. 140, No. 4, Pt. 2, October 2016

5th Joint Meeting ASA/ASJ   3443

5pSCb24. Temporal gap and amplitude modulation detection evaluated using the electrically evoked auditory change complex in patients with auditory brainstem implants. Shuman He, Tyler C. McFayden, Bahar S. Shahsavarani, and Katherine E. Gill (Ctr. for Hearing Res., Boys Town National Res. Hospital, 555 N. 30th St., Omaha, NE 68131, Shuman.He@boystown.org)

This study aims to objectively evaluate temporal processing capabilities using the electrically evoked auditory change complex (eACC) in patients with auditory brainstem implants (ABIs). Study participants include seven Cochlear Nucleus 24M ABI recipients (six children and one adult). All children received ABIs due to cochlear nerve deficiency and the adult recipient was diagnosed with neurofibromatosis type II (NF-2). The speech processor was bypassed and electrical stimulus was a biphasic pulse train directly sent through individual electrodes of the ABI. Gap durations tested in this study ranged from 4 to 256 ms. Amplitude modulation frequency was 10 Hz. Modulation depths tested in this study ranged from 1% to 99%. Behavioral gap detection thresholds were also obtained for two participants. The eACCs were evoked by temporal gaps and amplitude modulations. The shortest gap duration and the smallest amplitude modulation depth that can be used to evoke the eACC vary across study participants and across stimulating electrodes within individual participants. These preliminary data also suggest that the temporal gap detection primarily depends on subcortical auditory neural structures. The eACC can be used to objectively evaluate temporal processing capabilities in both adult and pediatric ABI recipients.

5pSCb25. An event related potentials method for assessing audiovisual speech integration: Implications for children with autism spectrum disorders. Julia Irwin, Lawrence Brancazio, Jacqueline Turcios, Trey Avery, and Nicole Landi (Haskins Labs., 300 George St., Ste. 900, New Haven, CT 06511, julia.irwin@yale.edu)

Perceptual studies of children with autism spectrum disorders (ASD) strongly implicate deficits in processing of audiovisual (AV) speech. Previous research with AV stimuli has typically been conducted in the context of auditory noise or with mismatched auditory and visual ("McGurk") stimuli. Although both types of stimuli are well-established methods for testing typically developing (TD) participants, they may create additional processing problems for children with ASD. To more precisely examine audiovisual (AV) speech perception in children with ASD, we developed a novel measure of AV processing that involves neither noise nor AV cross-category conflict. The speech stimuli include clear exemplars of the syllable /ba/ and a modified version of /ba/ in which the consonant is substantially weakened so that the syllable is heard as "/a/". These are dubbed with a video of the speaker saying /ba/. Audiovisual integration should result in the visual information effectively "restoring" the weakened auditory "/a/" cues so that the stimulus is perceived as /ba/. Using event related potentials (ERP), we will present evidence from typically developing adults and preliminary results from children with ASD and TD to examine whether children with ASD are weaker in AV speech integration.

5pSCb26. Naming practice app on tablets and communication robots for people with aphasia. Shingo Kuroiwa, Yasuo Horiuchi (Chiba Univ., 1-33, Yayoi-cho, Inage-ku, Chiba, Chiba 2600033, Japan, kuroiwa@faculty.chiba-u.jp), Sachiyo Muranishi, Daisuke Furukawa (Kimitsu Chuo Hospital, Kisarazu, Chiba, Japan), Kyohei Ishihata, Akihiko Morimoto (Robocure Inc., Tokyo, Japan), and Hiroji Suzuki (Escor Inc., Kisarazu, Japan)

Aphasia is a language disorder that negatively affects a person's ability to understand and generate language. Although it has been said that patients in the chronic phase are not likely to recover their ability, a number of recent studies report that highly frequent and continuous training enables recovery. Therefore, we have started a project that aims to improve the environment in which people with aphasia can continue language learning at home. In this presentation, we demonstrate naming practice apps, using speech recognition technology for people with aphasia. They can check whether their utterances are accurate, through the use of speech recognition and they can listen to the correct pronunciation of words through the use of speech synthesis. Field trials using the app for tablets have been conducted since Dec. 2014, and for the communication robot, Pepper, since Dec. 2015. Although the average of speech recognition accuracy was 61 % (false

acceptance: 2%, false rejection: 37%), the participants accepted the poor robustness of the speech recognition, and tried to speak clearly. Some participants felt as if the robot was somewhat similar to a child, and others understood its inflexibility, and so in consequence they both accepted its low speech recognition ability.

5pSCb27. Suprasegmental characteristics of spontaneous speech produced in good and challenging communicative conditions by younger and older adults. Outi Tuomainen and Valerie Hazan (Speech Hearing and Phonetic Sci., Univ. College London, 2 Wakefield St., London WC1N 1PF, United Kingdom, o.tuomainen@ucl.ac.uk)

Our study investigates strategies used to clarify our speech and compensate for masking effects when communicating in challenging listening conditions in older and younger adults. A total of 50 older (OA, 65-85 years, 30 F) and 23 younger adults (YA, 18-35 years, 14 F) were recorded (in "Talker A" role) while they completed the problem-solving diapix task with a young adult ("Talker B") in four listening conditions: with no interference (NORM), when Talker B had a simulated hearing loss (HLS), Talker B heard babble (BAB1) or both heard babble (BAB2) noise. We measured articulation rate, fundamental frequency (f0) median and range, energy in 1-3 kHz band reflecting spectral tilt (ME1-3 kHz) for Talker A. In NORM, OAs were slower speakers and had lower ME1-3 kHz and wider f0 range than YAs. Median f0 also converged for men and women in OA talkers. In adverse conditions, YAs slowed down their speech (HLS) and increased the f0 range (BAB1 and BAB2) more than OAs, and OAs raised their median f0 more than YAs (BAB2). These results indicate that there are age-related changes in suprasegmental characteristics of spontaneous speech, and that OA talkers use different strategies than YA talkers to clarify their speech.

5pSCb28. Fathers' and mothers' differential talk to sons and daughters with hearing loss. Mark VanDam, Carsen Jessup (Elson Floyd Coll. Med., Dept. Speech & Hear Sci., Washington State Univ., PO Box 1495, Spokane, WA 99202, mark.vandam@wsu.edu), and Tracy Tully (Com Sci Disord, Eastern Washington Univ., Spokane, WA)

In a natural family environment, fathers have been shown to talk less than mothers, and fathers may talk more to sons than to daughters. There is conflicting evidence showing mothers' word use increases talking both to sons and to daughters. Further, preschoolers who are hard-of-hearing have been shown to encounter a similar number of parent utterances as compared with their typically-developing peers. This work looks at fathers and mothers talking to their preschool boys and girls and explores whether child hearing loss affects parent word use. Here we used all-day recordings from a wearable child audio recorder, analyzed using automatic speech processing algorithms to determine estimated number of syllables spoken by fathers, mothers, boys, and girls. About half of the families included children with hearing loss. Results showed that mothers talked more than fathers, parents offered about the same amount of talk to boys or girls who were typically developing, fathers talked more to their boys who were hard-of-hearing, and mothers talked more their girls who were hard-of-hearing. The asymmetry of fathers talking more to boys with hearing loss and mothers talking more to girls with hearing loss is considered in terms of the differential experience hypothesis.

5pSCb29. Comparison of speech rate and intelligibility between children with cerebral palsy and typically developing children in Taiwan. Chu-Yin Wu, Li-mei Chen (Foreign Lang., National Cheng Kung Univ., 1 University Rd., Tainan 701, Taiwan, wu25689@gmail.com), Yu Ching Lin (Physical Medicine and Rehabilitation, National Cheng Kung Univ., Tainan, Taiwan), Katherine C. Hustad, and Raymond D. Kent (Waisman Ctr., Madison, WI)

Due to the defect of speech motor control, cerebral palsied children (CP) with dysarthria often show slower, weaker, and imprecise articulation with poor coordination, which have profound impact on speech intelligibility (Hodge & Gotzke, 2014; Hustad, Gorton, & Lee, 2010). It was reported that CP children had lower speech rate and intelligibility than typically developing (TD) children. However, little research discussed the application of this finding in languages other than English. The present study investigated the

correlations between speech rate and intelligibility in 4 Mandarin-speaking CP children aged 4-8 and 4 TD children aged 5. Speech data were collected from the repetition tasks of sentences ranging from 3 to 8 words. Speech rate were calculated through WPM (words per minute) and IWPM (intelligible words per minute), and speech intelligibility of each child was judged by three listeners. The results showed that: 1) TD had significantly higher mean intelligibility and slightly higher IWPM than the CP group; 2) Mean WPM of CP was higher than TD; 3) Higher WPM in CP did not correspond to higher intelligibility. Contrary to the previous findings, in the present study, CP showed higher WPM but lower intelligibility. This discrepancy requires further examinations in future studies.

5pSCb30. A comparative study of Mandarin compound vowels produced by prelingually deafened children with cochlear implants and normal hearing peers. Jing Yang (Commun. Sci. and Disord., Univ. of Central Arkansas, CSD/ Speech Lang. Hearing Ctr., 201 Donaghey Ave., UCA Box 4985, Conway, AR 72035, jyang@uca.edu) and Li Xu (Commun. Sci. and Disord., Ohio Univ., Athens, OH)

The present study examined the dynamic features of compound vowels in native Mandarin-speaking children with cochlear implants (CIs). Fourteen prelingually deafened children with CIs (aged 2.9−8.3 years) and 14 age-matched, normal-hearing (NH) children produced monosyllables containing six Mandarin vowels (/aɪ/, /aʊ/, /uo/, /iɛ/, /iaʊ/, /ioʊ/). The frequency values of the first two formants were measured at nine equidistant time points (10-20-30-40-50-60-70-80-90%) over the course of vowel duration. All formant frequency values were normalized and then used to calculate vowel trajectory length, vowel section length, overall spectral rate of change, and vowel section rate of change. The results revealed that the CI children produced significantly longer durations for all six compound vowels. They clearly showed positional deviation from the NH children on vowels /aɪ/, /uo/ and /iɛ/. The CI children's ability to produce formant movement for the compound vowels varied considerably. Some CI children produced relatively static formant trajectories for certain diphthongs whereas others produced certain vowels with greater formant movement than did the NH children. As a group, the CI children roughly followed the NH children on patterns of vowel dynamic spectral change, but they moved the articulators with a slower rate of change than did the NH children.

5pSCb31. Age estimation in Japanese speech based on feature selection. Atsushi Morimoto, Masahiro Niitsuma, and Yoichi Yamashita (Graduate School of Information Sci. and Eng., Ritsumeikan, 1-1-1 Nojihigashi, Kusatsu-shi 525-8577, Japan, morimoto-ASJ@slp.is.ritsumei.ac.jp)

This paper addresses age estimation of Japanese speech utterance, using paralinguistic information in speech. Our speech conveys not only linguistic information but also paralinguistic and non-linguistic information. Nonlinguistic information includes potentially valuable social information such as personal parameters, body conditions, gender and ages. This significant information, however, has not been investigated enough. The main purpose of our study was to develop less-invasive methods to extract this information, by proposing a method to classify speakers' age into five groups (20s, 30s, 40s, 50s, and 60s). 6375-dimentional audio features were extracted from 1579 samples of 913 male speakers extracted from CSJ (Corpus of Spontaneous Japanese) and feature selection was conducted based on Fisher's Ratio. Naïve Bayes with top 20 parameters of Fischer's ratio, yields highest accuracy rate of 51.2%.

5pSCb32. The articulation of direct bone conducted sound in noisy environment. Hidehiko Maeda (Dept. of Rehabilitation Sci., Health Sci. Univ. of Hokkaido, 1757, Kanazawa, Toubetsucho-Ishikarigun, Hokkaido 061-0293, Japan, maehide@hoku-iryo-u.ac.jp) and Kiyoshi Yonemoto (Dept. of Social welfare, Iwate Prefectural Univ., Takizawa, Iwate, Japan)

We have constructed the direct bone conducted sound (DBC) detection system which is connected with piezo electronic accelerometer and the titanium implant of bone-anchored hearing aid (Baha) implanted on the temporal bone. Our previous study has reported that the sound distortion of the DBC was less than that of conventional bone conducted sound (BC) due to its property of outputting the DBC without going through the skin. This

study reports the articulation of the DBC in noisy environment. The subject who output the DBC wears a Baha on his left side for hearing loss. He pronounced 50 monosyllables and words in quiet and extremely noisy environment. His voices were output simultaneously throughout three pathways, DBC, BC, and air conducted sound (AC). Sixteen participants with normal hearing heard those voices and answered how those were heard. In the quiet environment, the percentage of collect answers of the AC showed statistically significance (<.01) than those of the DBC and the BC. However, in extremely noisy environment, those of the DBC showed statistically significances (<.01) than those of the AC and the BC. This study clarified that the DBC kept the articulation without changing it in extremely noisy environment.

5pSCb33. The perceptual and acoustic characteristics of Japanese elderly speech—The relationship between cognition and speech motor performance. Takako Yoshimura, Makoto Kariyasu, Akie Saito, Minoru Toyama (Speech and hearing Sci. and Disord., Kyoto Gakuen Univ., 18, Gotanda-cho, Yamanouchi, Ukyo-ku, Kyoto 615-8577, Japan, takakoy@kyotogakuen.ac.jp), and Mariko Osaka (Ctr. for Information and Neural Networks, Osaka, Japan)

Recent studies showed the evidences that language complexity or cognitive load may deteriorate the speed and accuracy of speech. No study was found to explore the relationship between cognitive load and speech movement directly from the perspective of working memory (WM) capacity. In Reading Span Test (RST), WM capacity would be reduced for speech production as the number of sentences and target words increased. The aim of our study is to reveal perceptual and acoustic characteristics of the elderly speech in relation to WM capacity. It was hypothesized that slower speech rate and/or slurred speech would be observed for elderly adults with reduced WM capacity or low cognitive function. Twelve elderly adult speakers of Japanese participated in this study. The scores of Montreal Cognitive Assessment Japanese version was used to divide the participants into high and low cognitive groups. For RST, each participant was asked to read visually presented sentence(s) aloud and to recall target word(s). Reading performances were recorded by a microphone being set in front of each participant for later acoustic evaluations. The reading performances of two cognitive groups were compared and results regarding on the presence of slower speech rate and slurred speech, judged by experienced speech language pathologists, were presented. We will discuss the relationship between the cognitive load and speech movement in the elderly based on the acoustic evaluation.

5pSCb34. Effect of denture adhesives on enhancement pronunciation accuracy of denture wearers. Hyung Woo Park and Sang hwi Jee (IT, SoongSil Univ., 1212 Hyungham Eng. Bldg. 369 Snagdo-Ro, Dongjak-Gu, Seoul, Seoul 06978, South Korea, pphw@ssu.ac.kr)

People who have lost their natural teeth for any reason need to wear dentures (false teeth in other words) to eat and talk. And the teeth chewing role, and also has an important role in having the pronunciation. However, when people wear a denture, it causes some problems such as aesthetic or pronunciation. Among denture users, pronunciation errors may occur depending on how bonded and which location of false teeth inside of our mouth. In this paper, a comparative analysis on the bandwidth of formant frequency found in pronunciation errors is utilized in order to analyze pronunciation errors caused by different positions of dentures and thus evaluate the accuracy of pronunciation for each position. In the experiment result, that show the pronunciation accuracy enhancing around 5%, when using adhesives, who use denture correctly. And accuracy of pronunciation is evaluated differently depending on the position of dentures.

5pSCb35. The effects of deep brain stimulation on conversational and repeated speech. Diana V. Sidtis (New York Univ., 665 Broadway, Rm. 936, New York, NY 10012, diana.sidtis@nyu.edu) and John J. Sidtis (Geriatrics, Nathan Kline Inst., Orangeburg, NY)

The effects of the DBS on speech remain controversial. This approach used intelligibility by listeners and acoustic parameters as measures of conversation and repeated speech. Persons with DBS-STN provided

conversational speech samples and then repeated portions of their conversation in both on and off stimulation. Excerpts were arranged for a listening protocol adjusted for the Parkinsonian hypophonia. Dependent measures were intelligibility and difficulty ratings by 30 healthy listeners and acoustic measures of voice. Fewer words were correctly transcribed from conversation compared to repetition [$F(1,29) = 14.06$; $p = 0.001$]. Similarly, intelligibility for STN-DBS on worse than off [$F(1,29) = 9.462$; $p = 0.005$]. More words were correctly transcribed from repetition than from conversation with STN-DBS on [$t(29) = -2.342$; $p = 0.026$] and off [$t(29) = -3.619$; $p = 0.001$]. Comparisons of conversation versus repetition and DBS state revealed significant effects of both task and DBS on intelligibility. Repetition yielded better voice quality than conversation and DBS improved voice. DBS is mildly disruptive to conversational speech, but the effect of task is greater. The external example inherent in repetition may reduce the burden for speech, as is seen for gait and arm reach.

5pSCb36. Ideal binary masking and the integration of spectro-temporal glimpses in cochlear implant simulations. Vahid Montazeri, Shaikat Hossain, and Peter F. Assmann (School of Behavioral and Brain Sci., Univ. of Texas at Dallas, GR 41, Richardson, TX 75080, vahid.montazeri@utdallas.edu)

Ideal binary masking (IdBM) can improve listeners' speech recognition scores in the presence of background noise. It retains the time-frequency (T-F) regions where the target dominates the noise while discarding the remaining regions. Li and Loizou (J. Acoust. Soc. Am. **123**, 1673-1682, 2008) suggested that IdBM is effective because the pattern of target-dominated T-F units (IdBM pattern) directs the listener's attention to target glimpses. This study extends the above explanation and investigates listeners' abilities to integrate target glimpses when the IdBM pattern is held constant. Speech-shaped noise maskers were added to IEEE target sentences at a signal-to-noise ratio of -5 dB and the mixture was processed with IdBM. The IdBM-processed mixture was then decomposed into the IdBM-processed target and masker. The IdBM-processed target/IdBM-processed masker power ratio (TMR) was then varied to produce an altered IdBM-processed mixture. These processed signals were either sent to an eight-channel tone-vocoder or kept intact, and were presented to normal hearing listeners. Results showed that IdBM benefit heavily depends on TMR, especially in the vocoder condition. Overall, the results suggest that IdBM benefit is mediated by listeners' ability to integrate the target glimpses, despite the IdBM pattern being constant.

5pSCb37. Support system for pronunciation instruction and practice in special education classes for language-disabled children enrolled in conventional schools. Ikuyo Masuda-Katsuse (Kindai Univ., 11-6 kayanomori, Iizuka 820-8555, Japan, katsuse@fuk.kindai.ac.jp)

We developed two types of support systems for pronunciation instruction and practice in special education classes for language-disabled children: a stand-alone system and a web application system. The aims of the systems were (1) to encourage students to practice the pronunciation they learned in their classes repeatedly and (2) to promote cooperation between teachers and the speech-language-hearing therapists (STs) who support the teachers. The teachers registered practice words and their probable error pronunciations into the system, taking into account individual pronunciation needs, and the students were encouraged to read these practice words aloud. Three speech evaluation methods were prepared: automatic speech recognition, phonemic discrimination between the correct and probable error pronunciation of a consonant period, and articulation tests from STs. The web application system simplified access to the students' speech and exercise records for the STs. Practical field test results demonstrated that all of the teachers found that our system helped improve their students' pronunciation, the STs felt that it simplified not only articulation tests, but also observation of the students' pronunciation improvement process, and the STs and teachers believed that it would promote their mutual cooperation for more effective instruction.

5pSCb38. Levels of prominence in Spanish infant-directed speech. Ann M. Aly (Linguist, UCLA, 3125 Campbell Hall, Los Angeles, CA 90095, ann.m.aly@ucla.edu), Stephanie Jacobson (Palomar Health, Escondido, CA), and Megha Sundara (Linguist, UCLA, Los Angeles, CA)

Research on Spanish prosody in adult-directed speech (ADS) (Ortega-Llebaria & Prieto 2005, Ortega-Llebaria & Prieto 2011) shows evidence for three acoustic levels of prominence: stressed and focused, stressed and unfocused, and unstressed and unfocused. We tested whether Spanish infant-directed speech (IDS) contains the same levels of prosodic prominence as ADS. In order to test this hypothesis, 30-minute play sessions between adult Spanish-speaking females of Latin American origin ($n = 12$) and 12-month old infants were recorded. We extracted target words from the recordings containing stressed and unstressed /a/, /e/, /i/, and /o/ in non-final position. Voicesauce (Shue et al 2011) was then used to obtain acoustic correlates of stress in Spanish: pitch, duration, intensity. We also labeled pitch accents using Sp_{ToBI} conventions (Estebas Vilaplana & Prieto 2009). Latent class analyses to be performed on the data will determine the number of distinct prominence levels present in Spanish IDS as realized by Latin American speakers.

5pSCb39. Developmental progressions in the harmonic structure of infant-directed speech. Jhonelle Bailey (Marcus Autism Ctr., Children's Healthcare of Atlanta, Marcus Autism Ctr., 1920 Briarcliff Rd. NE, Atlanta, GA 30329, jbail22@emory.edu), Shweta Ghai, and Gordon Ramsay (Dept. of Pediatrics, Emory Univ. School of Medicine, Atlanta, GA)

Caregivers speak differently to infants than to adults early in life. Previous studies of this special register, "motherese," have sought to identify key differences in voice quality between infant-directed and adult-directed speech. Acoustic properties most salient to infants appear to reside in exaggerations of prosodic structure, but these findings are limited by speech analysis techniques that do not fully capture the acoustic structure of the maternal voice or the way motherese changes over the course of development. In this study, multitaper analysis was used to measure developmental progressions in the complete harmonic structure of the maternal voice. Samples of adult- and infant-directed speech were extracted from home audio recordings of 20 mothers collected monthly from 0-24 months using LENA technology. Multitaper analysis was used to calculate the time-varying amplitude and phase of every harmonic component, the residual noise component, and the spectral envelope, permitting a complete statistical analysis of source and resonance properties. Key differences and developmental changes between adult- and infant-directed registers were found not only in the fundamental frequency, but in the overall structure of the harmonic spectrum, indicating that manipulations of the whole voice source may be employed to create motherese.

5pSCb40. Acoustic correlates to word order in Infant- and adult directed speech: A cross-linguistic study. Irene de la Cruz-Pavía, Judit Gervain (CNRS-Université Paris Descartes, Laboratoire Psychologie de la Percept., 45, rue des Saints-Pères, Paris 75006, France, idelacruzpavia@gmail.com), Michael McAuliffe (McGill Univ., Montreal, QC, Canada), Eric Vatikiotis-Bateson, and Janet F. Werker (Univ. of Br. Columbia, Vancouver, BC, Canada)

The acoustic realization of phrasal prominence correlates systematically with the order of Verbs and Objects in natural languages. Prominence is realized as a durational contrast in V(erb)-O(bject) languages (English: short-**long**, *to [Ro]me*), and as a pitch/intensity contrast in O(bject)-V(erb) languages (Japanese: **high**-low, *['To]kyo ni*). Seven-month-old infants can use phrasal prominence to segment unknown artificial languages. This information might thus allow prelexical infants to learn the basic word order of their native language(s). The present study investigates whether this differing realization of phrasal prominence is also found in I(nfant) D(irected) S(peech), previously unexamined speech style. We recorded 15 adult native talkers of languages with opposite word orders producing target phrases

3446   J. Acoust. Soc. Am., Vol. 140, No. 4, Pt. 2, October 2016

5th Joint Meeting ASA/ASJ   3446

(English: VO, e.g., *behind furniture*, Japanese: OV, e.g., *jisho niwa*) embedded in carrier sentences, both in A(dult) D(irected) S(peech) and IDS, and conducted acoustic analysis of the phrases (i.e., pitch maximum, mean pitch and intensity, duration). The results revealed the expected contrast in pitch in Japanese IDS and a similar trend in ADS. In English, the predicted durational contrast was found in ADS, additionally accompanied by a pitch contrast. Interestingly, a pitch contrast but no durational contrast was found in IDS.

5pSCb41. Phonological development is affected by speech discrimination skills in typically developing children. Eija M. Aalto (Faculty of Humanistics/LogopeDC, Univ. of Oulu, Univ. of Oulu, P.O. Box 1000, Oulu 90014, Finland, eijaalto@gmail.com), Katri Saaristo-Helin, and Suvi Stolt (LogopeDC, Univ. of Helsinki, Helsinki, Finland)

In early language development accurate auditory discrimination is necessary for building phonological representations to mental lexicon. However, little is known about the connections between phonological development and auditory discrimination in noise in children at the age of 3;6 years. The aim of this study is to clarify the associations between phonological skills and auditory perception of 3;6 aged children (N = 65) acquiring Finnish. Phonological skills were assessed using Finnish Phonology Test (FPT) that has been normed previously. Auditory discrimination skills were assessed using a recognizing task which was created for the present study (ADT). In the task, children recognize and point from pictures (6-8 pictures in 5 subtasks) phonological neighbour words in babble noise (signal-to-noise ratio 13 dB). The results show that FPT and ADT scores correlated moderately (Spearman, r = 0.32, p = 0.01). Results support the view that auditory discrimination skills and phonological skills are associated in typically developed Finnish children. Thus, even typically developing children may benefit from decreased noise levels in their learning environment.

5pSCb42. Longitudinal profiling of speech development in bilingual infants at risk of autism. Shweta Ghai (Dept. of Pediatrics, Emory Univ. School of Medicine, 1920 Briarcliff Rd., Atlanta, GA 30329, shweta.ghai@emory.edu) and Jhonelle Bailey (Marcus Autism Ctr., Children's Healthcare of Atlanta, Atlanta, GA)

Bilingualism is common and increasingly prevalent worldwide, as are the autism spectrum disorders (ASD). However, there is little information on the impact of bilingual exposure on the development of language and communication skills in children with ASD. In the literature, only a few cross-sectional studies have analyzed language skills in both monolingual and bilingual toddlers with ASD, and have reported similar language outcomes using qualitative assessment instruments. Although outcomes may appear to be similar, it is not yet known if both groups follow the same developmental trajectories during infancy. To resolve this issue, this longitudinal study tracks speech development in monolingual and bilingual infants at low and high risk of autism using day-long audio recordings collected in the home with LENA technology from birth to 24 months. The relation between quantitative measures of speech development (adult volubility, child volubility, and conversational turns) over this period and qualitative language outcomes at 24 months is evaluated. The results of this study will help to determine the pace and patterns of language growth in bilingual infants relative to monolinguals at risk of autism, and thus provide insight into what may be alternative bases for the language acquisition process in bilinguals with ASD.

5pSCb43. Vowel final noise in child coda /s/. Helen M. Hanson (ECE Dept., Union College, 807 Union St., Schenectady, NY 12308, helen.hanson@alum.mit.edu), Stefanie Shattuck-Hufnagel (Res. Lab. of Electronics, Massachusetts Inst. of Technol., Cambridge, MA), and Jackson Doughty (ECE Dept., Union College, Schenectady, NY)

Child speakers of American English in the age range 2;6-3;6 often produce vowel final noise (VFN), or preaspiration, for [-voice] labial and velar stops in coda position of monosyllabic words [e.g., Shattuck-Hufnagel, Hanson, and Zhao, "Feature-cue-based processing of speech: A developmental perspective," in *Proc. ICPhS 2015*]. In contrast, the female caregivers of these children produce VFN less often than the children, with significantly

shorter durations. The current work extends the earlier work by studying a fricative, /s/, in coda position. Data from the same nine child/adult pairs are used. While the coda stops were all preceded by the vowel /ʌ/, the corpus includes two words with coda /s/, *Gus* and *goose*, allowing the observation of possible vowel effects. Preliminary analysis of a subset of the data shows that, as for coda stops, children are more likely than adults to produce VFN, with longer durations. However, adults show a vowel effect, being much more likely to produce VFN when /s/ follows /ʌ/ than when it follows /u/. A vowel effect is not observed for the child subjects.

5pSCb44. Cluster duration in a nonword repetition study of Russian-speaking children. Darya Kavitskaya (Slavic Lang. and Literatures, Univ. of California Berkeley, 6303 Dwinelle Hall, Berkeley, CA 94720, dkavitskaya@berkeley.edu) and Sharon Inkelas (Linguist, Univ. of California Berkeley, Berkeley, CA)

This paper reports results from a nonword repetition study of typically developing Russian-speaking children aged 7-10, finding that consonant cluster duration correlates with whether or not the cluster is lexically attested in Russian. Stimuli consisted of 144 nonwords, repeated once each. 36 nonwords began with biconsonantal (CC) clusters; 18 with triconsonantal (CCC) clusters. None of CCC clusters are attested in Russian. Half of the CC clusters are attested. The CCC clusters were frequently produced as CC, with deletion/omission of one of the three consonants. This study compares the durations of CCC clusters which were reduced to CC (example: stimulus *ptkoka*, produced as *tkoka*) to the CC clusters which were produced as intended. The main finding is that the mean duration of intended CC clusters (153 ms) is shorter than that of corresponding reduced (CCC to CC) clusters (215 ms) (n = 14, p < .01). Lexical attestedness also played a role. The mean duration of produced clusters which are lexically attested in Russian also tended to be shorter than that of lexically unattested clusters (159 ms vs. 214 ms). These findings suggest that both past production experience (the effect of attestedness) and planning complexity (the effect of reduction) affect cluster duration.

5pSCb45. Infant auditory and audiovisual speech discrimination. Kaylah Lalonde and Lynne A. Werner (Speech and Hearing Sci., Univ. of Washington, 1417 NE 42nd St., Seattle, WA 98105, klalonde@uw.edu)

Although infants are sensitive to the correspondence between auditory and visual speech cues, previous studies have not investigated whether infants can use this correspondence to improve speech perception in noise. This experiment used a modified observer-based psychoacoustic procedure to measure audiovisual benefit to speech discrimination in 6- to 8-month-old infants. Infants heard a single standard CV syllable repeatedly in a 65 dB A steady-state speech spectrum noise at 17 dB nHL (re: adult thresholds). Infants were trained to respond when they heard a new CV syllable. Some infants were tested with a vowel contrast; others were tested with a consonant contrast. Each infant completed either the auditory-only or audiovisual condition; when possible, they completed both conditions. In the auditory-only condition, a still image of the female talker remained on the screen throughout testing. In the audiovisual condition, a synchronous, congruent video of the talker played throughout testing. Preliminary results from 30 infants show a trend for better discrimination of audiovisual syllables than auditory-only syllables, suggesting visual speech improves infants' perception of speech in noise and could be useful in speech and language learning. Differences in results for vowel contrasts, consonant contrasts, between-subjects analysis, and within-subjects analysis will be discussed.

5pSCb46. The development of gendered /s/ perception in children and adolescents. Fangfang Li (Psych., Univ. of Lethbridge, 4401 University Dr., Lethbridge, AB T1K 3M4, Canada, fangfang.li@uleth.ca)

The sound /s/ in English is gender differentiating: women tend to produce a variant of /s/ with higher spectral mean frequency than men (Fox & Nissen, 2005; Stuart-Smith *et al.*, 2007). The gender marking cue inherent in /s/ frication is perceptually salient (Munson *et al.*, 2006; Strand & Johnson, 1996), but prior research has focused exclusively on adults. The current study aims to identify the developmental course of the perception of speaker gender when hearing isolated segments of fricative /s/. Sixty children and

adolescents from three age groups (7~8 years; 10-11 years; and 13-14 years) participated in the study. They were asked to listen to a series of /s/ segments and identity speakers' gender for each sound by clicking on a female's or a male's face on the screen. Our results indicate the developmental change in adolescents in the accuracy of gender perception and their association with the primary gender-marking cue in /s/ frication when comparing with the two younger groups of children. This study suggests that speech perception, particularly the type of perceptual knowledge related to socio-indexical aspects, can continue to develop beyond school years.

5pSCb47. Acoustic variation among African American and European American children: Age, gender, and ethnicity. Dalila Salas, Emily Simmons, Nicole Marsh, Julia Licata, and Sonja Trent-Brown (Psych., Hope College, 35 East 12th St., Holland, MI 49423, dalila.salas@hope.edu)

Previous research has shown vowels are acoustically specified based on their formant frequencies (Peterson & Barney, 1952). Hillenbrand *et al.* (1995) replicated and extended this study. While both included children, neither included variation across children's gender or ethnicity. The present study explores variations in phonemic production for children across age, gender, and ethnicity. Children were presented with a list of /h-vowel-d/ words and sentences ("*I hear the sound of /h-V-d/ some more*") containing 12 General American English vowels (/i,ɪ,e,ɛ,æ,ʌ,ɑ,ɔ,o,ʊ,u,ɝ/). Height and weight were also measured. We hypothesize that fundamental and formant frequencies will be higher for 8 to 9 year olds than for 10 to 12 year olds and higher for girls (by age 10). For ethnicity, we anticipate no significant differences in frequency measures between European-American and African-American children. This study contributes to our knowledge of developmental trajectories for specified acoustic parameters. As gender and ethnicity are vital cues for adult speakers, it is important to investigate how salient acoustic parameters are for child speakers and at what ages child parameters begin to approximate adult measures. Results will have implications for audiologists, speech language pathologists, developmental scientists, and others in the field of communication sciences and disorders.

5pSCb48. The role of auditory working memory In the development of glimpsing skills in children. Jessica R. Sullivan (Commun. Sci. and Professional Counseling, Univ. of West Georgia, 1601 Maple St., Carrolton, GA 30118, jreneesullivan@gmail.com), Erin C. Schafer (Univ. of North Texas, Denton, TX), and Homira Osman (Univ. of Washington, Seattle, WA)

Glimpsing is a complex process involving lower-and higher level auditory skills required for speech understanding in noise. However, little is know the role of working memory as glimpsing develops in children. The present study aimed to determine the extent to which auditory working memory contributes across the continuum of auditory perception skills in children. This study investigated the role working memory plays in speech recognition and auditory comprehension task in the presences of noise. We measured working memory, speech recognition and auditory comprehension performance with noise in 20 seven-to-nine year old children with normal hearing. Backward digit recall was used to assess working memory. Speech recognition in noise performance was measure using the Hearing In Noise Test. Comprehension was measured by a child's ability to answer five questions associated to auditorily-presented stories. Each question evaluated a different level of comprehension. In general, there are no relationships between the speech recognition in noise performance and working memory. However, the relationship between working memory and comprehension was stronger. These results suggest that working memory contributes more when the auditory task is more complex especially in the presence of noise.

5pSCb49. English-learning infants' perception of boundary tones. Megha Sundara (Dept. of Linguist, UCLA, 3125 Campbell Hall, Los Angeles, CA 90095-1543, megha.sundara@humnet.ucla.edu), Monika Molnar (Basque Ctr. on Cognition, Brain and Lang., Donostia, Spain), and Sonia Frota (Dept. of Linguist, Faculdade de Letras da Universidade de Lisboa, Lisbon, Portugal)

Like consonants and vowels, infants' ability to discriminate lexical pitch becomes language-specific with age (Mattock & Burnham, 2006). We know

less about when infants perceive pitch marking of prosodic units. We tested English-learning infants on Portuguese bisyllabic stimuli where statements and yes/no questions are segmentally identical but distinguished by a fall versus a rise on the final syllable. Using the visual habituation paradigm, we found that unlike Portuguese infants (Frota *et al.*, 2014), only English 8-, but not 4-month-olds are able to distinguish final rise from fall in the presence of segmental variability. Besides having a rising pitch, the final syllable in Portuguese questions is also longer. We are currently testing English 8-month-olds with resynthesized Portuguese stimuli that maintain the pitch difference, but neutralize duration. Together, the results will help us reconcile the previously reported failure of English-learning infants to distinguish English statements from questions (Soderstrom *et al.*, 2011; Geffen, 2014), with their precocious sensitivity to prosody (Jusczyk, 1997).

5pSCb50. Developmental trends of perceiving happy voice in preverbal infants. Feng-Ming Tsao, Yu-Hsin Hu, Chieh Kao (Psych., National Taiwan Univ., No. 1, Sec. 4, Roosevelt Rd., Taipei 106, Taiwan, tsaosph@mail2000.com.tw), and huei-mei liu (Special Education, National Taiwan Normal Univ., Taipei, Taiwan)

Perception of **emotional prosody** is essential for children to identify emotions of communication partners from utterances. Newborns prefer listening to happy talks but children start to weight emotional prosody over lexical meanings to judge emotions in speech around middle childhood. Happy voice is expressed with higher fundamental frequency (F0) and F0 variations in syllables manifest **lexical tones** in Mandarin. Would lexical tones affect happy prosody perception development in preverbal infants? The goal of this study was to explore whether the ability of perceiving happy prosody varies with lexical tones. Mandarin-learning children master the production of Tone 2 earlier than Tone 3, and development trend of lexical tone production might reflect perceptual trends of lexical tones. Thirty-nine Mandarin-learning 7- and 11-month-old infants were tested with two prosody (happy vs. neutral prosody) contrasts in a speech discrimination task. For each contrast, speech stimuli were monosyllabic words /ji/, varied with emotional prosody. Tonal information (Tone 2 or Tone 3) varied between contrasts and remained the same within a contrast. Results showed that older infants were more accurate than younger infants to distinguish happy prosody and the pitch contours of lexical tones did not greatly shift the developmental trend.

5pSCb51. Analysis of children's high front vowel area function using three-dimensional ultrasound imaging. Gary Yeung (Dept. of Elec. Eng., Univ. of California, Los Angeles, 420 Westwood Plaza, 56-125B Eng. IV Bldg., Los Angeles, CA 90095, garyyeung@g.ucla.edu), Steven M. Lulich (Dept. of Speech and Hearing Sci., Indiana Univ. Bloomington, Bloomington, IN), Asterios Toutios (Signal Anal. and Interpretation Lab., Univ. of Southern California, Los Angeles, CA), Abeer Alwan, and Amber Afshan (Dept. of Elec. Eng., Univ. of California, Los Angeles, Los Angeles, CA)

Understanding children's speech production can follow from an analysis of children's vocal tract area functions. While some studies have used imaging techniques on children to measure tongue contours, to our knowledge, no vocal tract area functions of children have been examined. We attempt to measure the oral cavity vocal tract area functions of children's vowels using three-dimensional ultrasound imaging recordings of the tongue and palate. In this investigation, we model the oral cavity area function of the vowel [i] of children and compare it to that of an adult oral cavity area function measured using the same ultrasound technique. We show evidence that the tongue positioning and articulation of the vowel [i] is dramatically different for children and adults. Specifically, we focus on how the different tongue positioning of children and adults affects the vocal tract transfer function of [i]. Some theories of children's speech, either mechanically or perceptually based, can be inferred from these findings and will be compared to previous studies. [Work supported in part by NSF Grant 1551113, NSF Grant 1551131, NSF Grant 1514544, and NIH Grant R01DC007124.]

5pSCb52. The time course of top-down processing during speech perception by normal hearing and cochlear-implant users. Anita E. Wagner (KNO, UMCG, Hanzeplein 1, Groningen 9713 GZ, Netherlands, a.wagner@umcg.nl)

Cochlear implant (CI) users need to rely on top-down processes, such as integration of semantic information, to facilitate perception of CI speech, degraded in acoustic-phonetic details. This paper investigates the time-course of semantic integration for normal-hearing (NH) listeners, when processing natural and CI-simulated speech, as well as for experienced CI users. In eye-tracking experiments, we recorded listeners' gaze fixations—an online measure of lexical decision-making, and time-locked pupil dilation—a measure of the mental effort involved. Participants' ocular responses were measured while listening to sentences and simultaneously performing a search among pictures that displayed objects referred to in the sentence. The stimuli contained a target word (baby) that was either preceded or followed by disambiguating context (crawl). The display showed next to the target also a phonological competitor (bay), a word that was semantically viable given the sentential context (worm), and an unrelated distractor. For natural speech, gaze fixations show a fast integration of sentential context, and pupil dilations reveal a release from mental effort when sematic information is integrated. Degraded signals delay semantic integration, and further constrain listeners' ability to compensate for inaudible information. We will discuss sources of individual variation among CI users.

5pSCb53. Aging effect on categorical perception of Mandarin level-rising and level-falling tones. Lilong Xu (Univ. of Cambridge, University of Cambridge, Cambridge CB21TN, United Kingdom, babybear747@hotmail.com), Xiaohu Yang (Speech, Lang., and hearing Ctr., Ctr. for Cross-Linguistic Processing and Cognition, Shanghai Jiao Tong Univ., Shanghai, China), Chang Liu (Dept. of Commun. Sci. and Disord., Univ. of Texas at Austin, Austin, TX), and Yuxia Wang (Speech, Lang., and hearing Ctr., Ctr. for Cross-Linguistic Processing and Cognition, School of Foreign Lang., Shanghai Jiao Tong Univ., Shanghai, China)

The primary purpose of the study was to investigate the aging effect on categorical perception of Mandarin Chinese tone 1 (level F0 pitch contour) and tone 2 (rising F0 pitch contour), as well as tone 1 and tone 4 (falling F0 pitch contour). Using level/rising and level/falling fundamental frequency continuum, both tone identification and discrimination was measured for older and younger native Mandarin listeners. Moreover, the stimuli duration was manipulated at 100, 200, and 400 ms with a purpose to examine whether longer duration made tone perception easier for the older listeners. Results reported reduced catgoricality in terms of shallower identification slope and smaller discrimination peakedness for the older listeners compared with their younger counterparts. Meanwhile, longer duration was observed with enhanced categorical perception of lexical tones for the older listeners. These results were interpreted as the decline of aging-related temporal/spectral processing and inhibitory control in the central auditory system.

FRIDAY AFTERNOON, 2 DECEMBER 2016   SOUTH PACIFIC 1, 1:00 P.M. TO 5:05 P.M.

Session 5pSP

Signal Processing in Acoustics: General Topics in Signal Processing III

Trevor W. Jerome, Chair

Department of Acoustics, The Pennsylvania State University, GTWT Water Tunnel, State College, PA 16804

Contributed Papers

1:00

5pSP1. Strong direct path interference suppression based on spatial filtering. Xueli Sheng (Underwater Acoust., Harbin Eng. Univ., No. 145 Nantong St., Shuisheng Bldg. 803, Harbin, China, shengxueli@aliyun.com), Xiaoyi Sun, Yali Shi, Longxiang Guo, and Dian Lu (Underwater Acoust., Harbin Eng. Univ., Harbin, Heilongjiang, China)

Strong direct path interference suppression is one of the problems to be solved in AUVs' cluster system which has the advantages of concealment, precision, anti-jamming capability and function distance. In order to estimate the target direction, 3 methods of beamforming, including covariance matrix taper(CMT) condition, null-constrained condition and second-order cone programing(SOCP) constraint condition, are used to restrain direct path interference suppression separately in this paper. Comparing the 3 methods with simulation, SOCP gives the weights taking main lobe response, sidelobe level, depth, and width of zeros into consideration; null-constrained condition has a smaller amount of calculate and better anti-interference ability. But when the array disturbance exists, all the methods' performances are not good. Then, an improved algorithm is proposed to improve the tolerance of array disturbances.

1:15

5pSP2. Comparison of the multiple maximum length sequences and multiple exponential sweeps method for parallel measurements of impulse responses. Shoma Kuroda (Graduate School of Eng., Oita Univ., 700 Dannoharu, Oita City 870-1192, Japan, v15e3010@oita-u.ac.jp), Shingo Uenohara, Keisuke Nishijima, and Ken'ichi Furuya (Faculty of Eng., Oita Univ., Oita, Japan)

The measurement of room impulse responses is a central problem in audio signal processing, particularly for spatial audio rendering and sound field reproduction applications. The calibration of modern rendering systems requires the knowledge of the room impulse responses between the loudspeakers and several possible listener positions, in order to compensate for the characteristics of the loudspeakers and the room. The parallel measurement of multiple loudspeakers can be efficiently performed using the multiple maximum length sequence (MMLS) or the multiple exponential sweep (MES) method. This study evaluates the MMLS and MES methods for performing parallel measurements of multi-channel acoustic systems. Moreover, we apply noise reduction to improve their measurement accuracies. Experiments were conducted to evaluate the accuracy, measurement duration, and computing time for the two methods. The results show that the MMLS method was as accurate as the MES method. Although the measurement duration was shorter for the MMLS method compared with the MES method, the computing time was longer.

5pSP3. Nonstationary signals, holograms, and concert halls. James B. Lee (None, 6016 S. E. Mitchell, Portland, OR 97206, cadwal@macforcego. com)

Sabine's analysis of sound in rooms involves a "mean-free-path," 4V/S, scaling linearly; for concert halls this measure is about 30 m, implying that perception in the time domain will be an issue. It is not possible to average signals over the space within a concert hall, for it is not possible to average speech and music over time. Speech and music are nonstationary signals, therefore concert halls cannot be ergodic systems. The necessary and sufficient condition for proper acoustics in large spaces is that every listener receive the same full information at some time. It is fulfilled if the space is a time-varying diffuse volumetric hologram with direct sound from the orchestral region acting as reference beam. If the space contains specularly reflected components those will be heard as echoes, representing information about size and shape of the hall; this "architectural information" develops at the expense of "musical information." Therefore, diffuse reflection is essential in the body of a concert hall.

1:45

5pSP4. Passive imaging of scatterers in an acoustic daylight imaging configuration. Jie Li, Peter Gerstoft (Marine Physical Lab., Scripps Inst. of Oceanogr., 9500 Gilman Dr. La Jolla, San Diego, CA 92093, jil004@ucsd. edu), Dazhi Gao (Ocean Univ. of China, Qingdao, Shandong, China), Guofu Li (Ocean Univ. of China, Tianjin, China), and Ning Wang (Ocean Univ. of China, Qingdao, China)

The arrival-time structure of Green's functions is extracted by crosscorrelating one-sided aeroacoustic surf noise in an acoustic daylight imaging configuration, where 14 microphones (separation 0.08-3.7 m) are located between the noise and a cylinder. The cylinder is a hollow PVC pipe (radius 20 cm, height 2 m) 2.5 m landside of a microphone array. The frequency band for computing the cross correlation function is 400-2000 Hz, and total averaging time is $T_{ave} = 3$ min. Results show that arrivals corresponding to the scattered (travel-time for source-scatterer-receiver path) waves and the time-difference (travel-time difference between two scatterer-receiver paths) waves emerge in the crosscorrelation functions for an acoustic daylight imaging configuration. Only one microphone pair is needed to locate the scatterer passively with the scattered and time-difference traveltime. We show imaging of 20 crosscorrelation functions with clear time-difference and scattered waves. Each estimated location is along the seaside boundary of cylinder, and the average location is close to the origin of the seaside half circle of cylinder.

2:00

5pSP5. Extracting elastic target information from acoustic scattering data of underwater munitions. David E. Malphurs (Naval Surface Warfare Ctr. Panama City Div., 110 Vernon Ave., Panama City, FL 32407, david. malphurs@navy.mil)

Synthetic Aperture Sonar (SAS) acoustic scattering data from a variety of underwater munitions was collected in the Target and Reverberation Experiment 2013 (TREX13), Bay Experiment 2014 (BayEX14) and from the Acoustic Test Facility at Naval Surface Warfare Center Panama City Division. One data product from these tests is two dimensional templates of back scattered acoustic intensity as a function of aspect angle and frequency. A common method of target classification utilizes normalized template cross-correlations as the input into classification algorithms. This is typically done in a manner agnostic to observable physics-based scattering phenomena derived from the physical size, material composition, and geometry of the targets that might help improve classification performance. This presentation discusses methods explored to process the available data to make target elastic phenomena easier to recognize so that associated information can be extracted and used as target classifier input. The methods explored include: lower dimensional representations of the data that preserve the physics-based phenomena, analysis of the data in time-frequency space, and using phase information in template correlations to help isolate resonance related information. [Work supported by ONR.]

2:15

5pSP6. A method for real-time, multiple-source localization using spherical harmonic decomposition. Jonathan Mathews and Jonas Braasch (Architectural Sci. - Architectural Acoust., Rensselaer Polytechnic Inst., 110 8th St, Greene Bldg., Troy, NY 12180, mathej4@rpi.edu)

By taking advantage of the spherical harmonic transform, spherical microphone arrays are able to decompose three-dimensional sound fields into their basic components, which enables detailed analysis and efficient spatial filtering algorithms. In recent years, methods for determining relative source positions around an array have been studied. Weighting functions can be generated to create a steerable beam which attenuates a sound field except in a specific direction. This research describes a real-time framework for multiple-source detection, which can be implemented using these methods. By finding the energy-sum distribution across a matrix of weighted harmonic signals, directional information may be simultaneously and efficiently obtained for multiple sources over time.

2:30

5pSP7. Multi-channel non-negative matrix factorization with binary mask initialization for automatic speech recognition. Iori Miura (Graduate School of Eng., Oita Univ., 700 Dannoharu, Oita, Oita 870-1192, Japan, v16e3019@oita-u.ac.jp), Yuuki Tachioka, Tomohiro Narita, Jun Ishii (Mitsubishi Electric, Kanagawa, Japan), Fuminori Yoshiyama (Graduate School of Eng., Oita Univ., Oita, Japan), Shingo Uenohara, and Ken'ichi Furuya (Faculty of Eng., Oita Univ., Oita, Japan)

Non-negative Matrix Factorization (NMF) factorizes a non-negative matrix into two non-negative matrices. In the field of acoustics, multichannel expansion has been proposed to consider spatial information for sound source separation. Conventional multi-channel NMF has a difficulty in an initial-value dependency of the separation performance due to local minima. This paper proposes initial value settings by using binary masking based sound source separation whose masks on the time frequency domain are calculated from the time difference of arrival of each source. The proposed method calculates initial spatial correlation matrices using separated sources by binary masking. The music separation experiments confirmed that the separation performance of the proposed method was better than that of the conventional method. In addition, we evaluated initial value settings by using binary masking for automatic speech recognition (ASR) tasks in noisy environments. The ASR experiments confirmed that appropriate initial value settings were effective because initializations starting from previously well-estimated spatial correlation matrices achieved better ASR performances than random initializations.

2:45

5pSP8. Performance improvement of snore activity detection in a noisy environment. Keisuke Nishijima, Shingo Uenohara, and Ken'ichi Furuya (Comput. Sci. and Intelligent Systems, Oita Univ., Dannoharu 700, Oita-shi, Oita 870-1192, Japan, k-nisijima@oita-u.ac.jp)

Health promotion and maintenance is becoming increasingly important and depends on three elements: nutrition, exercise, and rest (sleep). In the present work, focusing on sleep, we develop a smartphone-based system based on snore activity detection to investigate day-to-day variations in the sleep state, which does not require dedicated hardware. Snore activity detection is performed by machine learning using a support vector machine (SVM) with a linear kernel; the SVM is trained by labeled acoustic features, and the trained SVM models are used to detect snore activity. As acoustic features, The sound pressure level and mel-frequency cepstrum coefficients are calculated from sleep sound data obtained using a smartphone. In this paper, we investigated the effects of adding sleep environment noise recorded before sleep to the training set in snore activity detection, and we considered ways to improve detection performance. Performance comparison among the conventional method of SVM and the proposed method was presented. The performance comparison was evaluated by the cross-validation. The detection performance was improved by adding sleep environment noise recorded before sleep to the training set.

3:00–3:20 Break

3:20

5pSP9. MRI noise suppression using weighted spectral subtraction based on noise estimation with long-and-short terms. Sayaka Okayasu, Maiko Yoshiura, Takahiro Fukumori, Masato Nakayama, and Takanobu Nishiura (Ritsumeikan Univ., 1-1-1 Noji-higashi, Kusatsu, Shiga 525-8577, Japan, is0159ee@ed.ritsumei.ac.jp)

MRI devices are often utilized as surgery support systems. However, communication among medical staff during surgery is difficult because conversations can be disturbed by loud noise emitted from devices. In this study, we propose a MRI noise suppression method using a weighted spectral subtraction to address the problem. Spectral subtraction is generally utilized for suppressing stationary noise. However, directly applying this method to non-stationary noise such as MRI noise, which has frequency and amplitude fluctuations in the long-and-short terms, may be difficult. In the proposed method, the noise from a MRI device is estimated on the basis of the frequency fluctuation in the short term. The estimated noise is then subtracted by processing the weighted spectral subtraction whose coefficients are calculated on the basis of the amplitude fluctuation in the long term. Objective evaluation experiments were carried out to evaluate the quantities of noise reduction and sound distortion. As a result, we confirmed the effectiveness of the proposed method.

3:35

5pSP10. Feature extraction for acoustic signatures of small boats. Alexander S. Pollara, Alexander Sutin, and Hady Salloum (Maritime Security Ctr., Stevens Inst. of Technol., 1 Castle Point on Hudson, Babbio Ctr., Hoboken, NJ 07030, apollara@stevens.edu)

Small boats represent a particular challenge to maritime surveillance due to their ubiquity, low radar cross-section, and absence of AIS transmission. We present a review of previously developed acoustic classification methods and demonstrate the importance of both the engine, and propeller as the two sources of small vessel noise. The features used for classification are extracted using various signal processing methods including signal spectra, spectra of Detection of Envelope Modulation on Noise (DEMON), and Cepstrum. The normalization of the features allowing to remove the dependence of the parameters on the distance to a boat and its speed was conducted. Acoustic signatures of six small boats including a panga, a jetski, an electrical boat, a go fast boat and two speed boats of the type common in the United States were collected by Stevens in a large glacial lake in NJ. The sound sources of various tonal components in the signal spectra are discussed. The program for automated extraction of classification features is presented and the identification of each of the 6 small boats with these features is demonstrated. [This work was supported by DHS's S&T Directorate.]

3:50

5pSP11. Numerical simulations of sound source localization with two-dimensional bio-inspired antennas of varying geometric complexities. Michael Reinwald (Laboratoire d'Imagerie Biomédicale, Université Pierre et Marie Curie, 15 rue de l'Ecole de Médecine, Laboratoire d'Imagerie Biomédicale, Paris 75006, France, mchlrnwld@gmail.com), Lapo Boschi (Institut des Sci. de la Terre de Paris, Université Pierre et Marie Curie, Paris, France), Stefan Catheline (LabTAU, Unité U1032, Université Claude Bernard Lyon 1, Lyon, France), and Quentin Grimal (Laboratoire d'Imagerie Biomédicale, Université Pierre et Marie Curie, Paris, France)

Many animals use audition as their primary tool to navigate and hunt. The accuracy with which they complete those tasks suggests that they effectively harness detailed waveform information. In addition to inter-aural time (ITD) or level differences (ILD) they are likely to extract information from sound conducted by the skull. It is well established that an object of complex shape, in which elastic waves propagate (such as the skull of an animal), can be used as an antenna to locate the position of an acoustic source, even using a single receiver. We hypothesize that complexities in the shape of mammalian skulls enhance source localization performance. We evaluate the localization precision achieved with two-dimensional skull-shaped antennas of varying complexity, generating synthetic data by time-domain numerical simulations and "inverting" them via signal processing techniques such as ITD, ILD and time-reversal. Full waveform algorithms (e.g., time-reversal) benefit from increasing geometric complexity of the antenna, e.g., removing

symmetry ambiguities in the localization. We investigate this relation for aforementioned algorithms and combinations thereof. Our simulations are a first step towards experimental and numerical investigations of the role of bone conduction in source localization using skulls of different mammals.

4:05

5pSP12. Applying the unit circle constraint to the diagonal loaded minimum variance distortionless response beamformer. Colin Ryan (Elec. and Comput. Eng., Univ. of Massachusetts Dartmouth, 679 Main St., Harwich, MA 02645, cryan2@umassd.edu) and John R. Buck (Elec. and Comput. Eng., Univ. of Massachusetts Dartmouth, Dartmouth, MA)

Estimating the direction of arrival of a narrowband planewave using a uniform line array is a common array processing problem. Capon's minimum variance distortionless response (MVDR) adaptive beamformer (ABF) suppresses interferers by steering nulls in their directions while maintaining unity gain in the desired look direction. Practical ABFs replace the ensemble covariance matrix (ECM) with a sample covariance matrix (SCM) estimated from array observations, or snapshots. When the ABF cannot average enough snapshots, the SCM is a poor estimate of the ECM, and may be rank deficient. Applying diagonal loading (DL) to the SCM improves the ABF's performance by limiting the white noise power, but may also limit interferer suppression. Projecting the zeros of the beamformer's array polynomial onto the unit circle (UC) provides deeper nulls to suppress interferers but exhibits worse white noise behavior than DL. The UCDL MVDR beamformer combines the UC constraint with the DL MVDR weights to achieve better SINR than either the UC or DL MVDR alone. Even when the UCDL beamformer suffers mismatch on the DL level, the UCDL SINR rivals or betters the DL MVDR with the optimal DL level. [Research funded by ONR grant N00014-15-1-2238.]

4:20

5pSP13. Sound source localization based on sparse estimation and convex clustering. Tomoya Tachikawa, Kohei Yatabe, Yusuke Ikeda, and Yasuhiro Oikawa (intermedia art and Sci., Waseda Univ., 59-407-2,3-4-1, Okubo, Shinjuku, Tokyo 169-8555, Japan, 14919320tt611@toki.waseda.jp)

Sound source localization techniques in array signal processing have been interested for many years. A lot of direction-of-arrival (DOA) estimation methods have been proposed. Most of them use plane waves as model. On the other hand, 3D sound source localization techniques are important for many applications. They are required to estimate the DOA and the distance of sound sources. In the case of the techniques targeted to estimate the distances of sound sources, monopoles must be used for the model. The positions of monopoles are candidates for the sound source positions. However, it is difficult to handle a large number of monopoles placed at various distances. Moreover, it is difficult to decide the number of sound sources to be estimated from the candidates. The proposed method uses sparse estimation with a monopole dictionary in frequency domain, as is well known that the effectiveness of sparse representation for DOA estimation has been in late years. After the sparse estimation, the proposed method uses convex clustering to estimate the DOA and the distance of sound sources. In addition, the proposed method includes the idea that decides the number of all sound sources without a priori information about the number of sound sources.

4:35

5pSP14. Context recognition system for smartphone applications with acoustic data. Miho Tateishi (School of Sci. for Open and Environment Systems, Keio Univ. Graduate School of Sci. and Technol., Hiyoshi 3-14-1, Kohoku-ku, Yokohama 223-8522, Japan, m.tateishi-3303@keio.jp)

This paper aims to create an original context recognition system on smartphones by using acoustic data. In this paper, the context represents that how the smartphone users move or how their contiguous environment is. In current systems, the context refers to the place where they are. It means that it is difficult to recognize the context when they are in a new place for which the system does not have the data. Our system simply categorizes the situation: "Train" when the user on the train, "Quiet Work Place" when they in a place like a library, PC room, or laboratory, and so on. This system features sound because it would be containing a lot of information. From volume to spectrum, we analyze the data from a built-in microphone and abstract multiple feature values. In addition to this,

accelerator and luminance data are also used as feature values. Then these feature values are classified into context categories. This situation based categorizing system realizes that wherever the user goes, the context can be understood. It would be a flexible platform for smartphone applications which need the context recognition system.

5pSP15. Identification of phonocardiographic signals using noise suppression based on wavelet transform. Shinya Kudo, Keisuke Nishijima, Shingo Uenohara, and Ken'ichi Furuya (Comput. Sci. and Intelligent Systems, Oita Univ., 700 Dannnoharu, Oita 870-1192, Japan, v15e3009@oita-u.ac.jp)

While heart disease is one of the three major diseases, only well-qualified doctors can evaluate phonocardiographic signals. However, phonocardiographic signals are not always used in healthcare because only a few professionals are experts in evaluating phonocardiographic signals. We need to develop an easily available system that can automatically diagnose phonocardiographic signals. In previous study, phonocardiographic signals are typically analyzed using wavelet transform to match and extract the characteristics of known normal and abnormal phonocardiographic signals. However, everyday noises such as lung and breath sounds, environmental noises, and blood flow noises may contaminate these signals and hinder analysis. We have previously proposed noise suppression using wavelet transform for phonocardiographic signals. In this study, we compare the proposed method and spectral subtraction to identify the kind of phonocardiographic signals. The experiment results show that the proposed method provides better identification performance compared with the spectral subtraction method.