Session 5aAB

Animal Bioacoustics: General Topics in Animal Bioacoustics II

Ted W. Cranford, Chair
Dept. of Biology, San Diego State Univ., 5500 Campanile Dr., San Diego, CA 92182

Contributed Papers

8:15

5aAB1. Hearing threshold measurements using auditory evoked potentials of four stranded short-finned pilot whales (*Globicephala macrorhynchus*) in Key Largo, FL. Danielle R. Greenhow (College of Marine Sci. Univ. of South Florida, 140 7th Ave. S, Saint Petersburg, FL 33701), Micah Brodsky (Dolphins Plus, Key Largo, FL 33037), Robert Lingenfelser (Marine Mammal Conservancy, Key Largo, FL 33037), and David A. Mann (College of Marine Sci., Univ. of South Florida, St. Petersburg, FL 33701)

Approximately 26 short-finned pilot whales, *Globicephala macrorhynchus*, stranded in Cudjoe Key, FL, on May 5, 2011. Four animals, two adult and two juvenile females, were transported to a rehabilitation facility in Tarpon Basin, FL. Auditory evoked potentials (AEPs) were recorded in response to amplitude modulated tone pips modulated at 1000 Hz. AEP thresholds were determined at 10, 20, 40, 80, and 120 kHz for all the four animals. Audiograms were similar to previous findings in pilot whale hearing tests. Short-finned pilot whales have a lower peak sensitivity than other odontocetes such as bottlenose dolphins. Greatest sensitivity was 40 kHz for all whales, while thresholds for the two adults were 25–61 dB higher at 80 kHz than those for the juveniles tested. Click evoked potentials were similar between the four whales and comparable to other echolocating odontocetes. Five total pilot whales have been tested during two separate stranding events; the previously tested juvenile male was found to have profound hearing loss (Mann *et al*., 2010; Schlundt *et al*., 2011). These findings add to the limited database of pilot whale (short- and long-finned) hearing studies, of which there are only two others (Schlundt *et al*., 2011; Pacini *et al*., 2010).

8:30

5aAB2. Monaural and binaural aerial hearing thresholds in a California sea lion (*Zalophus californianus*). Arial M. Brewer (G2 Software Systems Inc., 4250 Pacific Hwy., Ste. 125, San Diego, CA 92110), Jason Mulsow, and James J. Finneran (Space and Naval Warfare Systems Ctr. Pacific, Code 71510, San Diego, CA 92152)

Though studies have focused on some aspects of binaural hearing in otariid pinnipeds, none have directly measured monaural hearing thresholds. Recent auditory evoked potential (AEP) testing with a young male California sea lion suggested that hearing threshold differences potentially existed between the subject's two ears. To further investigate these findings, aerial psychophysical audiograms were collected for this sea lion. Using headphones to deliver stimuli, monaural and binaural thresholds were measured at seven frequencies from 0.5 to 32 kHz in one-octave steps. The binaural and two monaural (left and right ears) psychophysical audiograms were all similar in that they had the typical mammalian U-shape, with peak sensitivity from 8 to 16 kHz. The right ear showed generally elevated psychophysical thresholds when compared to the left, although all differences were less than 7 dB. The psychophysical and AEP audiograms were similar in terms of shape; however, the directions of monaural threshold differences determined with AEP methods were not always consistent with those obtained using psychophysical methods. [Work supported by ONR.]

8:45

5aAB3. Hyperoodon echo-morphology: Biosonar anatomy in the northern bottlenose whale. Ted W. Cranford (Biology Dept., San Diego State Univ., 2674 Russmar Dr., San Diego, CA 92123-3422), Marianne H. Rasmussen (Univ. of Iceland, 640 Hsavik, Iceland), Charles W. Potter (Natl. Museum of Natural History, Washington, DC, 20560), and Petr Krysl (Univ. of California at San Diego, La Jolla, CA, 92093)

The northern bottlenose whale (*Hyperoodon ampullatus*) is one of the largest toothed whales, marked by an outsized bulbous forehead. Like other odontocetes, the head contains a sophisticated biosonar system. The *in situ* anatomic geometry of the biosonar components is normally inaccessible by the large size of these animals. We acquired a 400 kg head from a 6.18 m female Icelandic bottlenose whale. The head was encased in a special sarcophagus. An industrial CT scanner generated more than 700 scans at 2 mm thick, from which we reconstructed the detailed anatomy of the head. The acoustic fats are ensheathed in higher density (acoustically reflective) tissues that act to channel sounds into and out of the head. Lipid channels begin at both sets of phonic lips, coalesce into a single S-shaped channel that eventually passes between two large maxillary crests, and then forms a horn shaped melon that projects sound into the environment. Sounds are apparently received over the jaws and throat and travel back to the bony ear complexes through lipid pathways. These lipid channels bifurcate and insert onto the ear complex where the bone is thin. The finite element simulations for biosonar sound transmission and will be reported.

9:00

5aAB4. Anatomy of a Northeastern Pacific Ocean Sciaenid Fish, the White Weakfish (*Atractoscion nobilis*). Carl R. Schilt (Bigleaf Sci. Services, P.O. Box 225, North Bonneville, WA 98639) and Ted W. Cranford (San Diego State Univ.)

Anatomy of a Northeastern Pacific Ocean Sciaenid Fish, the White Weakfish (*Atractoscion nobilis*) Carl Schilt, Bigleaf Science Services, North Bonneville, WA and Ted Cranford, San Diego State University, San Diego, CA. The bony fish family Sciaenidae, commonly called the croakers, grunts, and drums, are as bioacoustically interesting as their names suggest. Although the anatomy and bioacoustics of some North Atlantic Ocean forms have been investigated, those of the Pacific Ocean are less-well studied. The white weakfish (or white seabass, *Atractoscion nobilis*) is a near-shore marine sciaenid of the West Coast of North America and is the target of sport and commercial fisheries and the subject of a substantial stock enhancement program. Two fresh post-mortem specimens of the white weakfish were subjected to CT scanning on two different spatial scales. The full body of the larger animal, about 1 m long total length (TL), was scanned at a 0.625 mm voxel size whereas the head of a much smaller specimen (about 20 cm TL) was scanned in a "micro-CT scanner with a voxel size of 90 μm. Images and measurements of both specimens, especially bioacoustically relevant anatomy and comparisons with other sciaenid species, will be presented.

9:15

5aAB5. Performance of an echolocating bottlenose dolphin in the presence of anthropogenic masking noise. Eryn M. Wezensky (Sensors and SONAR Dept., Naval Undersea Warfare Ctr., Code 1554, 1176 Howell St., Newport, RI 02841, emwezens@nps.edu), James J. Finneran, Jason Mulsow (US Navy Marine Mammal Program, SSC Pacific, San Diego, CA 92152), Cameron R. Martin (Naval Res. Enterprise Internship Program, Ste. 600, Washington, DC 20036), Brian K. Branstetter, Patrick W. Moore, and Dorian S. Houser (Natl. Marine Mammal Foundation, 2240 Shelter Island Dr. #200, San Diego, CA 92106)

To study the effects of acoustic masking from anthropogenic noise, bottlenose dolphin (*Tursiops truncatus*) echolocation performance was assessed in the presence of different masking noise types using Navy relevant source transmissions. Echolocation clicks produced by the dolphin were detected with a hydrophone, then digitized within a phantom echo generator (PEG). The PEG converted the received clicks into echoes, delayed appropriately for the simulated target range. The echoes were then broadcast to the dolphin via a sound projector while masking noise transmissions were held constant. Using an acoustic response and a modified method of constants procedure, the echolocation performance of the dolphin was computed as a function of range between 3 and 17 m. Comparative echolocation performance to different masking noise type categories was analyzed between intermittent and continuous noise, direct path transmissions and multipath exposure, and mid frequency versus high frequency bands. These results expand the limited understanding of biosonar processing capability and signal characteristic alterations used to discriminate and resolve changes in small scale features while exposed to potential noise interference types. [Work supported through the Office of Naval Research.]

9:30

5aAB6. Dolphins (*Tursiops truncatus*) use multiple auditory mechanisms for detecting tonal signals in natural and anthropogenic noise. Brian K. Branstetter (Natl. Marine Mammal Foundation, 2240 Shelter Island Dr., #204), Jennifer S. Trickey, Kimberly L. Bakhtiari (G2 Software Systems Inc., San Diego, CA 92110), and James J. Finneran (US Navy Marine Mammal Program)

The dolphin's ability to detect a 10 kHz tone masked by a variety of noise types was measured using a standard band-widening paradigm. Maskers included natural noise (rain, snapping shrimp, and ice squeaks), anthropogenic noise (pile saw and boat propeller cavitation) and statistical noise (Gaussian and comodulated noise). For most noise types, detection thresholds increased as noise bandwidth increased up to 1 kHz (the dolphin's critical bandwidth at 10 kHz). Masking patterns for narrow-band maskers were similar regardless of the masker type. However, for noise bandwidths greater than 1 kHz, masking patterns diverged by as much as 23 dB depending on the noise type. The power spectrum model provided reasonable predictions for Gaussian, rain, pile saw, and boat noise masking patterns. Additional experiments suggested that mechanisms related to temporal envelope processing (across-channel envelope comparison and within-valley listening) determined masking patterns for snapping shrimp and comodulated noise. Thresholds in ice squeak noise, which proved to be the most effective masker, were related to the dolphin's inability to discriminate the signal from the background noise rather than inability to detect the signal. These results suggest that the dolphin auditory system uses multiple mechanisms for signal detection in complex noise [Work supported by the ONR.]

9:45

5aAB7. Auditory effects of multiple underwater impulses on bottlenose dolphins (*Tursiops truncatus*). James J. Finneran (US Navy Marine Mammal Program, SSC Pacific Code 71510, San Diego, CA 92152, james.finneran@navy.mil), Jennifer S. Trickey (G2 Software Systems Inc., San Diego, CA 92110), Brian K. Branstetter (Natl. Marine Mammal Foundation, 2240 Shelter Island Dr., #204 San Diego, CA 92106), Carolyn E. Schlundt (ITT Corp., San Diego, CA 92110), and Keith Jenkins (US Navy Marine Mammal Program, San Diego, CA 92152)

It is becoming increasingly clear that noise may adversely affect the hearing and behavior of many marine mammals, including dolphins and whales. One of the possible consequences of noise exposure is a temporary loss of hearing sensitivity, called a temporary threshold shift (TTS). Previous studies of TTS in marine mammals have used only steady-state and single impulsive sounds; however, many naval, seismic survey, and construction activities generate multiple underwater impulsive sounds. To investigate the auditory effects of multiple underwater impulses, auditory thresholds were measured in bottlenose dolphins before and after exposure to a series of impulses produced by a seismic airgun. Hearing thresholds were compared to determine the amount of TTS as a function of exposure level and the number of impulses. Thresholds were measured by using both psychophysical and auditory evoked potential methods. Preliminary results reveal no measurable TTS after exposure to 10 impulses, with a total cumulative sound exposure level of approximately 195 dB re 1 μPa^2 s. [Work supported by the EP Sound and Marine Life Joint Industry Programme.]

10:00

5aAB8. Subject-based calibration of a jawphone transducer with a bottlenose dolphin (*Tursiops truncatus*). Jason Mulsow (U.S. Navy Marine Mammal Program, Space and Naval Warfare Systems Ctr. Pacific, Code 71510, 53560 Hull St., San Diego, CA 92152), Arial M. Brewer (G2 Software Systems Inc., 4250 Pacific Hwy., Ste. 125, San Diego, CA 92110), and James J. Finneran (U.S. Navy Marine Mammal Program, Space and Naval Warfare Systems Ctr. Pacific, San Diego, CA 92152)

Hearing studies with odontocete cetaceans often use suction-cup transducers known as''jawphones" to imitate underwater stimulus transmission. Jawphones are typically calibrated by measuring a frequency-response curve with a receiving hydrophone placed at a controlled distance in an underwater direct field. This procedure is somewhat controversial, as it may not sufficiently reproduce the odontocete sound reception pathway. This study calibrated a jawphone by comparing two behavioral audiograms for a single bottlenose dolphin. The first audiogram comprised underwater hearing thresholds (in dB re 1 μPa) measured using direct-field stimulation, the second comprised thresholds (in dBV) measured using a jawphone. All thresholds were measured using a psychophysical staircase procedure at frequencies from 14.1 to 150 kHz. The calibration curve was calculated by subtracting, at matched frequencies, jawphone thresholds from underwater direct-field thresholds. The resulting curve had the shape of a bandpass filter, with highest levels at frequencies from 56.6 to 130 kHz. This subject-based curve was similar in shape to a previously obtained direct-field calibration curve, although the subject-based values were higher above 20 kHz. These results are especially relevant to auditory evoked potential hearing studies that measure thresholds in odontocetes that are untrained for psychophysical procedures. [Work supported by ONR.]

10:15–10:30 Break

10:30

5aAB9. Migrating gray whale vocalizations and concurrent visual observations near Santa Barbara, California. Lauren K. Roche (Scripps Inst. of Oceanogr., UCSD, 9500 Gilman Dr., Mail Code 0205, La Jolla, CA 92093-0205, lroche@ucsd.edu), Michael H. Smith (Gray Whales Count, CA, 93101), Ana Širović (Scripps Inst. of Oceanogr., CA 92093-0205), Allison M. Cope (Gray Whales Count, CA 93101), Jasmine S. Buccowich, Sean M. Wiggins, and John A. Hildebrand (Scripps Inst. of Oceanogr., CA 92093-0205)

The vocalizations of migrating eastern North Pacific gray whales *(Eschrichtius robustus)* are not well understood, but studying these sounds may provide insights to their behaviors and encounters along their migratory route. To record migratory gray whale sounds over long periods, a high frequency acoustic recording package (HARP) was deployed from March to May 2010 off the north coast of the Santa Barbara Channel, an area where the gray whale northbound migration is close to shore. The HARP recorded continuously and concurrently with a shore-based visual survey of marine mammals. Comparisons of calls and sightings were conducted to relate calling repertoires with various types of migrating whale behaviors and groups (e.g., cow/calf pairs, juveniles, etc.). The spatial and temporal overlap of the visual and acoustic data enabled the estimation of call detection ranges and source levels. Characterization of gray whale vocalizations also will help with the development of automatic detectors which will aid in future the investigations of long-term gray whale recordings, including potential responses to anthropogenic activity.

10:45

5aAB10. Characterization of shape diversity in bat biosonar beampatterns. Mohammad Motamedi, Lvyin Cai (Dept. of Mech. Eng., Virginia Tech, IALR, Danville, VA 24540), Qiuping Xu, Washington Mio (Dept. of Mathematics, Florida State Univ., Tallahassee, FL 32306), and Rolf Müller (Dept. of Mech. Eng., Virginia Tech, IALR, Danville, VA 24540)

Numerical beampattern estimates for a large digital shape data set with 361 three-dimensional models of the pinna and noseleaf shapes from 106 different bat species are currently in preparation. Shape samples in this data set have been selected to represent taxonomic diversity and balance noseleaves versus pinnae. The data set will be used to analyze the biodiversity in bat beampatterns. In order to facilitate the extraction of interspecific patterns, spherical harmonics are being evaluated as an analysis tool to describe the spatial frequency content of the beampatterns. Spherical harmonics are the angular portion of a set of solutions to Laplace's equation represented in a system of spherical coordinates. The higher the harmonic, the more solutions are in the respective set. Regardless of the number of solutions, each set forms an orthogonal system. For a compressed characterization of the beampatterns, the representation of the beampattern shape at a given frequency is expressed as a power spectrum; the norm of the vector of the weighting coefficients for all basis functions in the set of a harmonic is taken to represent the portion of the total power explained by a particular harmonic. This power spectrum was found to reflect several beampattern properties well.

11:00

5aAB11. Non-rigid pinna deformations in bats—Behavioral adaptation or dynamic sensing?. Rolf Müller (Dept. of Mech. Eng., Virginia Tech, IALR, 150 Slayton Ave., Danville, VA 24540, rolf.mueller@vt.edu), Li Gao, and Lin Feng (School of Phys., Shandong Univ. 250100 Jinan, China)

Mammal groups with exceptional spatial hearing abilities, bats and primates in particular, are known to have intricate pinna shapes. These shapes determine the monaural spatial hearing sensitivity as a function of direction and frequency, i.e., the reception beampattern. Since even comparatively small, local pinna features can have profound effects on the beampattern, it is possible that some of the interspecific diversity in these pinnae represents spatial hearing adaptations on an evolutionary time scale. In addition, some bat groups, such as horseshoe bats, have elaborate muscular actuation mechanisms that can produce non-rigid pinna deformations. These mechanisms could enable beampattern control on a much shorter, behavioral time scale. This could add considerably to the capabilities of the biosonar systems in these species, because by switching between different pinna shapes—and hence reception beampatterns—could meet even conflicting demands of different biosonar tasks. However, since pinna deformations in horseshoe bats can happen on comparatively short time scales (i.e., one tenth of a second), it is also conceivable that the bats' pinnae are in motion during the reception of an echo. It is not clear at present if and how such a dynamic receiver paradigm could be utilized to enhance biosonar performance.

11:15

5aAB12. Reception beamforming in horseshoe bats is a dynamic process. Lin Feng (School of Phys., Shandong Univ., Shanda South Rd. 27, Jinan 250100, China, lfeng@sdu.edu.cn) and Rolf Müller (Dept. of Mech. Eng., Virginia Tech, IALR, Danville, VA 24540)

Horseshoe bats execute conspicuous pinna movements as part of their biosonar behaviors. These movements contain rigid rotations as well as deformation components that can result in considerable changes to the pinna shapes. Since pinna shape determines the biosonar system's sensitivity as a function of direction and frequency, these non-rigid deformations could be functionally relevant. In the work reported here, it has been investigated how the non-rigid deformations of pinna shape relate to the arrival of echoes. For this purpose, multi-camera high-speed video recordings of both pinna of a bat were synchronized with ultrasonic recordings of the emitted pulses and the returning echoes from a microphone placed next to the head of the animals. Landmark dots were painted on the pinna to enhance the determination of position and velocity from stereo pairs of video recordings. It was found that the pinnae were deformed between as well as during echo receptions. Hence, the echoes were received by different static each pinna shapes (e.g., bent or upright) as well as by pinna shapes that underwent non-rigid deformation within the duration of the echo. As a consequence, echo reception by horseshoe bats involves dynamic beampatterns that can change between as well as during echoes.

11:30

5aAB13. Characterization of dynamic baffles in biosonar and biomimetic devices. Sajjad Z. Meymand, Mittu Pannala, and Rolf Müller (Dept. of Mech. Eng., Virginia Tech, IALR, 150 Slayton Ave., Danville, VA 24540, sajjadzm@vt.edu)

The pinna shapes of horseshoe bats are actively actuated to undergo non-rigid deformations on a time scale of one tenth of a second. These cycle times are similar to the durations of the echoes that the animals receive. Hence, it is possible that the bat pinnae are undergoing a change in shape while an individual echo is impinging on them. Such a dynamic reception paradigm based on a continuously deforming baffle are an interesting concept for explaining the biosonar as well as the design of bioinspired devices. Beampatterns are the standard characterization of time-invariant reception devices. In a beampattern, device gain is given as a function of direction and frequency. The Fourier transform that underlies this representation does not represent time variant behavior. Hence, new ways to characterize the time-variant behavior of a deforming baffle have to be found. This is further complicated because a complete description of system behavior requires four independent variables: azimuth, elevation, and either two time variables or time and frequency. Representations that are either time-frequency or entirely time based and use different ways to visualize the overall system behavior have been implemented and tested on the responses of biomimetic pinna shapes.

11:45

5aAB14. Time-variant biomimetic beamforming baffle prototypes augmented with local shape features. Mittu Pannala, Ojili Praveen Kumar Reddy, Sajjad Z. Meymand, and Rolf Müller (Dept. of Mech. Eng., Virginia Tech, IALR, 150 Slayton Ave., Danville, VA 24540, mpannala@vt.edu)

Horseshoe bats are known to exhibit non-rigid pinna deformations as part of their biosonar behaviors. During these deformation cycles, the pinnae of the animals alternate between upright and bent shape configurations that differ in their overall geometry as well as in the shapes and relative positions of local shape features. Artificial baffle prototypes have been designed to emulate these deforming bat pinnae. The prototype shapes were designed by augmenting an obliquely truncated cone with biomimetic local shape features. The prototype was implemented using elastic (synthetic rubber) materials. The deformation of the structure was accomplished through a simple actuation mechanism that was inserted on the back side of the artificial pinna. Features of the horseshoe bat pinna that have been evaluated for this study include a prominent vertical ridge that runs along the entire inner pinna surface, the antitragus along with the lateral incisions that separate it from the back portion of the pinna, and a pinna-edge incision near the tip of the pinna. An experimental analysis of deforming prototypes has demonstrated that such biomimetic local shape features affect the beampattern of the static artificial pinna shape as well as the time-varying properties of the deforming prototype.

Session 5aNSa

Noise and Physical Acoustics: Launch and Aircraft Noise

Victor W. Sparrow, Chair
Graduate Program in Acoustics, Pennsylvania State Univ., 201 Applied Science Bldg., University Park, PA 16802

Chair's Introduction—8:00

Contributed Papers

8:05

5aNSa1. A semi-analytical ray method to predict the propagation of long-range vertical noise: Application of NORD 2000 to the prediction of aircraft en-route noise. Kieran Poulain and Victor Sparrow (Penn State, Grad. Program in Acoust., 201 Appl. Sci. Bldg., University Park, PA 16802)

In the past, considerable efforts were made into developing models for the prediction of short-range propagation, in the context of airport noise (Integrated Noise Model, FAA) or more generally for short-range horizontal community noise propagation (NORD 2000, Delta Inc.). Recent research efforts have focused on the prediction of aircraft en-route noise for flights above 18 000 feet above ground level (AGL) (5.49 km AGL), in order to estimate the noise impact in U.S. National Parks and other quiet areas. Unlike other community noise issues, long-range vertical propagation requires an altitude-stratified and realistic atmosphere, which directly impacts the geometrical and absorption losses as well as the ground effects. Given the large distances involved, the use of a semi-analytical propagation model based on the ray theory could prove to be useful in reducing the computation time. NORD 2000, based on semi-analytical ray theory, was modified for use with aircraft en-route noise. A comparative study against AERNOM (Advanced En Route NOise Module, based on the numerical ray method) from Penn State is presented. [Work supported by VOLPE National Transportation Systems Center. The findings are the views of the authors and do not necessarily reflect the views of VOLPE, the FAA, NASA, or Transport Canada.]

8:20

5aNSa2. Evaluation of procedural changes for minimizing noise impact near military airfields. Andrew Christian, Victor Sparrow (Graduate Program in Acoust., The Penn State Univ., 201 Appl. Sci. Bldg., University Park, PA 16802, azc144@psu.edu), Micah Downing, and Kevin Bradley (Blue Ridge Res. and Consulting, Asheville, NC 28801)

Over time, communities surrounding military airfields tend to grow, while the noise generated by subsequent generations of military aircraft tends to increase. This creates an accelerating problem of human exposure to noise. One technique to alleviate this disparity is to modify attributes of flight trajectories in a way that reduces the sound exposure of the population. This involves the changing of flight ground tracks, profiles, and power settings within linear and non-linear boundaries which can include exceptions (such as no-fly-zones.) To demonstrate the effectiveness of these changes, DNL levels are calculated for real world population distributions and existing Navy departure procedures. These values are generated computationally using the widely accepted NoiseMap model. It is shown that significant reductions of exposure averaged over the population can be achieved in this way. Possible methods for automating this procedure are discussed. Properties of the objective function spaces are explored (e.g., convexity) and the appropriateness of different optimization methods is discussed. [Work supported by Naval Air Warfare Center.]

8:35

5aNSa3. Similarity spectra and the effect of propagation on F-22 spectra at afterburner. Tracianne B. Neilsen and Kent L. Gee (Dept. of Phys. and Astronomy, Brigham Young Univ., N283 ESC, Provo, UT 84602)

The afterburner spectra of a static F-22 Raptor have been compared to the similarity spectra from the two-source theory for jet noise. In the two-source theory, fine-scale turbulent structure (FSS) account for sideline radiation and large-scale structures (LSS) dominate the radiation for inlet angles aft of 125 deg for model-scale jets [Tam and Zaman, AIAA J. **38**(4), 592–599 (2000)]. A comparison of the similarity spectra with F-22 afterburner data, recorded at ground-based microphones along the 40-ft foul line, shows good agreement between the FSS spectrum and the sideline radiation and the LSS spectrum and the spectrum measured at 150 deg. However, discrepancies exist between the measured spectrum in the direction of maximum radiation and the LSS spectrum. Most significantly, the high frequency roll-off in the data is shallower (approx. 20 dB/decade) than predicted by the LSS (29.8 dB/decade). Most likely this difference comes from nonlinear propagation of the noise, which could then cause the agreement with the similarity spectra to be dependent on distance. Results from the ongoing investigation into how the measured spectra evolve relative to the FSS and LSS spectra as the noise propagates are presented. [Work sponsored by the Air Force Research Laboratory.]

8:50

5aNSa4. A parabolic equation method for predicting sound propagation above and below a layered ground. Hongdan Tao and Kai MIng Li (Dept. of Mech. Eng., Purdue Univ., 140 S. Martin Jischke Dr., West Lafayette, IN 47907-2031)

A numerical model based on a wide-angle parabolic equation (PE) technique has been developed for calculating the sound fields above and below a layered ground. A finite element discretization was applied along the vertical direction instead of the classical finite difference scheme. By using the finite element approach, the boundary conditions, i.e., the continuity of pressure and velocity, can be incorporated directly at the air/ground and ground/ground interfaces. The range-dependent sound fields were obtained by marching the finite-element solutions (above and below the layered ground) in the radial direction. This paper reports a preliminary formulation and presents some initial computational results for the sound fields above and below a porous ground. The numerical results from the PE calculations were compared with the results obtained from analytical solutions for and other benchmark results. A linear and cubic Hermite interpolate function was used in the numerical formulations for the finite element model. It has been shown that use of the cubic Hermite interpolation function generally leads to more accurate numerical solutions with fewer elements. [Work partially funded by Federal Aviation Administration and the China Research Scholarship Council.]

9:05

5aNSa5. Shock associated noise generation in curved turbulent Coanda wall jets. Caroline Lubert (Dept. of Mathematics & Statistics, James Madison Univ., Harrisonburg, VA 22807)

Curved three-dimensional turbulent Coanda wall jets are present in a multitude of natural and engineering applications. The mechanism by which they form a shock-cell structure is poorly understood, as is the accompanying shock-associated noise (SAN) generation. This paper will discuss these phenomena from both a modeling and experimental perspective. The Method of Characteristics will be used to rewrite the governing hyperbolic partial differential equations as ordinary differential equations, which will then be solved numerically using the Euler predictor-corrector method. The effect of complicating factors such as radial expansion and streamline curvature on the shock-cell location predictions will be discussed. Comparison of the theoretical calculations of the shock wave structure with associated schlieren flow visualization results will be presented. Related acoustical measurements will also be communicated. In this way, critical flow characteristics for shock-cell formation will be identified, and their influence on SAN discussed.

FRIDAY MORNING, 4 NOVEMBER 2011 ROYAL PALM 1/2, 9:45 TO 11:50 A.M.

Session 5aNSb

Noise: General Topics in Noise

Ann E. Bowles, Chair
Hubbs Sea World Research Inst., 2595 Ingraham St., San Diego, CA 92109

Chair's Introduction—9:45

Contributed Papers

9:50

5aNSb1. Vuvuzelas and their impact. Kenneth A. Cunefare (The Georgia Inst. of Technol., School of Mech. Eng., Atlanta, GA 30332, ken.cunefare@me.gatech.edu), Richard Ruhala, Laura Ruhala (Div. of Eng., Southern Polytechnic St. Univ., Marietta, GA 30060), and Tina Ortkiese (Arpeggio Acoust. Consulting, LLC, Hampton, GA 30228)

Vuvuzelas are inexpensive plastic horns that became a part of the American vocabulary during the 2010 World Cup held in South Africa, though horns of the style have been employed for decades in the States. During the World Cup, spectators played these horns through the games, causing noise complaints from athletes and spectators alike, including television listeners. The vuvuzelas and similar small, plastic horns are sold in the U.S. and other countries, and there is concern that they will create a related noise problem at future sporting events. This work presents the sound pressure levels, spectra, and directivity measured in a hemianechoic chamber produced by several different vuvuzelas, one at a time. Furthermore, a model is developed to predict noise levels within the stands and on the playing field assuming type- and temporal-mixes of the horns.

10:05

5aNSb2. Theoretical and numerical modeling of a parallel-baffle rectangular duct. Davide Borelli, Corrado Schenone, and Ilaria Pittaluga (DIPTEM, Univ. of Genova, Via all'Opera Pia 15/A, I 16145, Genova, Italy)

The paper describes the theoretical and numerical analysis of sound attenuation in a parallel-baffle rectangular ducts. Insertion losses in a frequency range up to 8000 Hz were predicted by means of a FEM numerical model and by means of analytical models from Sabine and Kurze. The models were then validated in the frequency range from 125 to 8000 Hz by comparing theoretical and numerical results with experimental data obtained in accordance to EN ISO 11691 and EN ISO 7235 standards. The results of the comparison indicate that the behavior of such a dissipative/reactive silences, with its internal-reflections and energy dissipation phenomena, can be predicted quite well by the FEM model on the whole frequency range. On the contrary, analytical models show little accuracy and such predictions are not always so accurate as design requires; besides, the complexity of the analytical approach tends to limit its application to the common design practice. Overall comparisons suggest that FEM modeling can be an accurate and inexpensive way to predict sound attenuation in parallel-baffle mufflers and fulfile the ever rising needs of proper methods in acoustic design of AC and ventilation plants.

10:20

5aNSb3. Active control of a centrifugal fan in a mock laptop enclosure. J. James Esplin, John K. Boyle, Scott D. Sommerfeldt, and Kent L. Gee (Dept. of Phys. and Astronomy, Brigham Young Univ., N283 ESC, Provo, UT 84602, jamesesplin@hotmail.com)

Active noise control (ANC) has shown promise in minimizing the effect of fan noise on users. Recent research by the authors has developed a model which is used to implement ANC on the inlets of centrifugal cooling fans. This model is based on minimizing radiated acoustic power in a model of the fan radiation and using those results to determine appropriate nearfield locations for the error sensor(s). Though this approach has been experimentally verified in an idealized setting, it was not verified in a more realistic situation. This paper describes how this model was expanded from its idealized setting to a mock laptop enclosure. When necessary modifications to the model were made, tonal noise can be predicted in the nearfield of the fan inlets, which allows one to develop an effective compact, realistic ANC setup for use in the mock laptop enclosure. With this ANC setup, significant global reduction of the inlet tonal noise can be achieved.

10:35

5aNSb4. A two-dimensional model for control of centrifugal fan inlet noise in a notebook computer. John K. Boyle, J. James Esplin, Scott D. Sommerfeldt, and Kent L. Gee (Dept. of Phys. and Astron., Brigham Young Univ., N283 ESC, Provo, UT 84602, jkb321@gmail.com)

Previous work on active control of exhaust noise from small centrifugal fans demonstrated significant reductions of the blade passage frequency (BPF) tone. A fan and heat-sink were placed within a mock-up notebook

computer case, and control of the fan exhaust noise was measured. It was found that control of the BPF in the exhaust did not significantly affect noise radiated from the fan inlets into the notebook casing, suggesting that exhaust noise and inlet noise may be controlled separately without one adversely affecting the other. In the current work, a two dimensional half-space, source coupling model has been developed to calculate the field within the notebook casing caused by the inlet noise. As a first approximation, free-space boundary conditions were used. A two-dimensional space was constructed to test the model, and error sensor placement was predicted. Measurements of radiated sound power show significant reduction of the blade passage frequency tone. Factors influencing experimental agreement with the model are discussed, such as modal effects and primary source location.

10:50

5aNSb5. Noise reduction in pipelines using Helmholtz resonators. Maaz Farooqui, Samir Mekid, Muhammad Hawwa, Atif Aliuddin, and AbduAziz Al-Hamoud (Mech. Eng., King Fahd Univ. of Petroleum and Minerals, Dhahran 31261, Saudi Arabia, maazfarooqui@kfupm.edu.sa)

Compressors are at the heart of most petrochemical and industrial power plants. They are usually the source of noise generation transmitted to pipelines. This noise is undesirable for people living close to installation and can also potentially cause structural failures in the piping. Particular attention is given to the modified solution using the Helmholtz resonator concept to reduce noise. An assessment of the noise reduction results using various types of resonators is carried out. A comparison between analytical, numerical simulation using COMSOL and experimental test is presented and discussed using a known patch-test.

11:05

5aNSb6. Compliant-wall methods for fluid-borne noise reduction. Kenneth A. Marek, Nicholas E. Earnhart, and Kenneth A. Cunefare (Georgia Tech Dept. of Mech. Eng., 771 Ferst Dr. Grad Box 334, Atlanta, GA 30332, ken.marek@gatech.edu)

One means of addressing fluid-borne noise in hydraulic systems is to add a compliant-walled section to the flow path. The impedance mismatch at the section boundary reflects a substantial amount of acoustic energy back to the source, and additional damping may be present due to the compliant material. While hoses can be used to add this compliance, the hose wall vibration produces undesirable breakout noise from the system. Additionally, the requirement that the hoses contain system pressure determines the minimum practical hose stiffness. An alternative approach is to use a silencer composed of a rigid outer housing, an internal compliant liner, and a central flow path through the liner. Such a configuration would avoid the pressure bearing and breakout noise complications while still adding compliance to the system. These two approaches are compared in this study, with simulated transmission loss presented for various lengths and material properties of the compliant sections.

11:20

5aNSb7. Sound quality characteristics for transient noise of high speed railway interior. Buhm Park (Dept. of Mech. Eng., Hanyang Univ., Haengdang-dong, Seongdong-gu, Seoul, Korea, parkbuhm@gmail.com), Sinyeob Lee (Hanyang Univ., Seoul, Korea), Sunghoon Choi (High-Speed Rail Div. Korea Railroad Res. Inst., Uiwang-si, Kyungki-do, Korea), and Junhong Park (Hanyang Univ., Seoul, Korea)

This study presents interior noise characteristics of Korean high-speed train under different operation conditions. The interior noise was measured for various operating conditions such as different speeds, concrete and ballast track, and open and tunneled track, passing by another moving train. The transient variation of the sound pressure was recorded. The objective and subjective sound quality evaluation of the interior noise was performed. For quasi-static sounds measured for open lands and in tunnel, Zwicker's loudness has dominant impact on the annoyance. When there are transient variations in the sound pressure encountered during entrance, exit, and passing by, the sharpness and roughness also have significant impact on the perceived annoyance. The obtained data are useful for the design of interior of the train since the transient variation of the sound pressure level is influenced also by the sound environment in the room.

11:35

5aNSb8. Measurement and analysis of squeal noise of vehicle wiper system. Dongki Min (Dept. of Mech. Eng., Hanyang Univ., 17 Haengdang-dong, Seongdong-gu, Seoul 133-791, Korea, dkmin@hanyang.ac.kr), Seongbin Jeong (Hanyang Univ., Seoul 133-791, Korea), Heewon Kang (Res. & Development Div. Hyundai Motor Co.), Honghee Yoo (Hanyang Univ., Seoul 133-791, Korea), and Junhong Park (Hanyang Univ., Seoul 133-791, Korea)

The squeal noise occurring from rubber and hard surface is an important issue, especially for vehicle wiper system. The noise is generated at relatively high frequency ranges, and its generation mechanism during the wiper operation is not straight forward to understand. In this study, experimental setup to generate the squeal noise in a consistent manner is designed using a rotating glass table. The wiper is located on the glass table in compressed status so that to simulate the wiper in the actual operating system. While the glass was rotating with constant velocity, the windshield washer fluid was sprayed on the glass. Under the normal constant velocity of the wiper operation, the condition (temperature, moisture contents, velocity) of continuous occurrence of the squeal noise was identified. The friction coefficient was measured, and its impact on the squeal noise generation was identified. The influence of various parameters such as the compression ratio, temperature, types of the wiper, angle of wiper installation, kinds of washer liquids was investigated using the suggested experimental setup. This provide the required information to design the vehicle wiper system that does not induce the squeal noise during the normal operation.

Session 5aSCa

Speech Communication: Speech Rhythm in Production, Perception and Acquisition II

Amalia Arvaniti, Chair
Dept. of Linguistics, Univ. of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0108

Chair's Introduction—8:00

Invited Papers

8:05

5aSCa1. Is infants' native language preference in the visual modality guided by speech rhythm?. Nancy Ward and Megha Sundara (Dept. of Linguist., UCLA, 3125 Campbell Hall, Box 951543, Los Angeles, CA 90069-1543, nancyward@ucla.edu)

Newborns prefer to listen to their native language. We have shown that this preference for the native language also exists with visual cues alone (Ward and Sundara, 2010). Four and 8-month-old monolingual English-learning infants looked longer at a face speaking English compared to Spanish. What information in the talking faces facilitates language identification, and thus language preference in infants? Studies on visual language discrimination (Soto-Faraco *et al*., 2007; Ronquest *et al*., 2010) suggest three possibilities: (1) lexical information accessed through lip-reading, (2) language-specific movement of the articulators, and (3) rhythmic information. Infants in their first year of life cannot rely on their developing lexicon for visual language preference. Here, we investigate whether rhythmic information is guiding infants' attentional preferences. We use the preferential looking procedure to test whether monolingual English and monolingual Spanish-learning 4-month-olds show a preference for looking at a face speaking Dutch over Spanish. Monolingual Spanish-learning 4-month-olds are expected to show a native language preference for the face talking in Spanish, replicating our previous findings. Crucially, if familiarity with language rhythm is sufficient to guide preferences for talking faces, we expect monolingual English-learning infants to demonstrate a preference for rhythmically similar Dutch, when presented opposite a rhythmically different Spanish.

8:25

5aSCa2. A cross-linguistic investigation of word segmentation by monolingual infants learning Spanish and English. Megha Sundara and Victoria Mateu (UCLA, Dept. of Linguist., 3125 Campbell Hall, Los Angeles, CA 90095)

Infants' ability to find words in fluent speech is affected by their language experience. The focus of this study was the extent to which monolingual Spanish- and monolingual English-learning 8-month-old infants' segmentation abilities are tied to their familiarity with the rhythmic properties of their native language. Spanish and English are considered to belong to different rhythm classes; the former is syllable-timed, whereas the latter is stress-timed. However, in both languages CVC.CV words are predominantly trochaic—i.e., words with stress on the initial syllable (Cutler Carter, 1984; Pons Bosch, 2010). We tested segmentation of trochaic words of the form CVC.CV in Spanish and English. Spanish word segmentation was tested using the Headturn Preference Procedure and English word segmentation using the visual fixation procedure. According to Jusczyk *et al*.'s rhythm hypothesis, infants should fail to segment words in a rhythmically different language. Contrary to the rhythm hypothesis, preliminary data indicate that Spanish-learning infants are successful in segmenting trochees in both Spanish and English. Thus, successful transfer of word segmentation abilities by infants seems to depend upon a match in the prosodic shape of the target word across the two languages.

8:45

5aSCa3. The learnability of language-specific fundamental frequency-based rhythmic patterns. Kristine M. Yu (Dept. of Linguist., Univ. of Mass. Amherst, 226 South College Bldg., Amherst, MA 01003, krisyu@linguist.umass.edu)

In addition to durational patterns, melodic patterns from variations in f0 contribute to percepts of speech rhythm, too (June 2005, Niebuhr 2009). The overall regularity of these melodic patterns is language-specific. In lexical tonal languages like Mandarin, f0 patterns may be quite variable, as many or all syllables may be associated with a tone; in stress-accent languages like English, f0 variation may be less variable as it is driven by stress and prosodic boundaries (Eady, 1982). In accentual phrase languages like Bengali and French, f0 variation is typically quite regular as it is driven almost entirely by prosodic boundaries in a very constrained way. With only acoustic parameters from the input available, are melodic patterns from these different kinds of languages learnable? And how is learnability affected if speech is directed to young infants? To study this, we recorded parents reading language samples in laboratory speech and in simulated infant directed speech in the languages mentioned above, and we are analyzing the speech with durational and f0 variability measures and prosodic (ToBI) labels. We will present results on learnability of melodic patterns using computational methods for pattern classification from a selection of the languages.

9:05

5aSCa4. The acquisition of English rhythm as a function of changes in phrase-level prosody. Melissa A. Redford, Hema Sirsa, and Irina A. Shport (Dept. of Linguist., Univ. of Oregon, Eugene, OR 97403-1290, redford@uoregon.edu)

Language rhythm emerges from a combination of lexical and phrase-level prosody. This talk will review evidence from on-going and completed studies on the acquisition of English prosody in typically developing school-age children to show that the surprisingly protracted acquisition of English rhythm is due to the extended acquisition of phrase-level prosody rather than to immature lexical stress production. Cross-sectional and longitudinal data from 20 native American-English speaking children indicate that whereas adult-like rhythm may not be fully acquired until age 8, even the youngest children have mastered lexical stress. Unlike eight-year-olds, though, younger children show less reduced vowels in function words and a less robust pattern of phrase-final lengthening. We argue, based on evidence from a separate study on stress-shifting in 25 six-year-olds and 25 adults, that younger children have yet to fully integrate lexical patterns into phrase-level prosodic structures. A lack of complete integration between lexical and phrasal levels may account for the extended acquisition of prosodically conditioned vowel reduction, which in turn could account for the prolonged acquisition of English rhythm. [Work supported by NIH Grant R01 HD061458.]

9:25

5aSCa5. Weighting of prosodic cues in language discrimination by infants and adults. Chad J. Vicenik and Megha Sundara (Dept. of Linguist., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095)

Previous research has shown that infants and adults can discriminate between prosodically similar languages using only prosodic cues. These experiments were designed to determine whether listeners use pitch cues or segmental duration and timing cues (i.e., rhythm cues) in language discrimination. We tested American English learning 7-month-old infants and adults on their ability to discriminate between sentences in American English and German that had been re-synthesized to isolate different cues. Infants were presented with low-pass filtered speech and speech with re-synthesized sinusoidal pitch. Adults were presented with low-pass filtered speech and two conditions where sonorants were replaced with /a/ and obstruents were replaced with silence: one contained both pitch and durational cues and the other contained only durational cues. Pitch cues were important for infants as well as adults. For infants, pitch cues were necessary for successful language discrimination; neither segmental nor durational cues were sufficient. For adults, the presence of pitch prevented listeners from discriminating English and German. Only when pitch cues were removed did adults rely on durational differences to distinguish languages. These results demonstrate that to distinguish prosodically similar languages, specifically English and German, American English listeners' weight pitch cues over durational cues.

9:45

5aSCa6. Rhythm, tempo, and F0 in language discrimination. Tara Rodriquez and Amalia Arvaniti (Dept. of Linguist., UC San Diego, 9500 Gilman Dr. #0108, La Jolla, CA 92093-0108, aarvaniti@ucsd.edu)

Existing literature on language rhythm suggests that languages can be discriminated in AAX "oddball" experiments only if they belong to different rhythm classes. Since exceptions to this pattern have been noted (e.g., the discrimination of Polish from English though both are said to be stress-timed), a series of five AAX experiments was run to test whether successful discrimination depends on factors other than rhythm class. English utterances were used as the base (AA) with Polish, Greek, Korean, Spanish, and Danish as the target languages. The stimuli were converted to sasasa while either retaining their original F0 or with flat F0; further, either the original tempo of each stimulus was retained or the stimuli (within each experiment) were manipulated to all have the same tempo (2 * 2 overall design, tempo * F0). The results show that tempo played the major role in discrimination with F0 cues being used when tempo differences between the two languages were eliminated. Rhythm class, on the other hand, played no significant part in discrimination suggesting that previous experiments may have interpreted as an effect of rhythm class the fact that tempo is often (though not always) slower in languages classified as stress-timed.

10:05–10:20 Break

Contributed Papers

10:20

5aSCa7. Quantifying rhythm: Interspeaker variation in %V. Margaret E L Renwick (Dept. of Linguist., Cornell Univ., Ithaca, NY 14850, mer56@cornell.edu)

The long-standing intuition that languages differ systematically in speech rhythm received support from an acoustic measure, %V (the ratio of vocalic material to the total duration of an utterance), which seemed to separate languages according to their perceived rhythms. Across studies, however, computations of %V values for target languages differ, raising the possibility that %V reflects not language-specific rhythmic habits, but the syllable structure of the particular utterances selected for analysis, and individual speaker differences. This paper replicates the %V methodology (using new recordings), and seeks to explain the high variability in %V by explicitly testing the link between %V, phonotactics, and individual speaker behavior. A high positive correlation is found between %V and an utterance's proportion of open syllables, while %V correlates negatively with the number of complex onsets in an utterance; these results support the hypothesis that utterance-level variation in syllable structure contributes substantially to the computed value of %V. However, the strength of this relationship varies across speakers. The goal of this paper is to determine the relative effects of utterance-level phonotactics, and individual speaker variation, on %V results, e.g., by comparing results when the same sentence is spoken by different speakers of a single language.

10:35

5aSCa8. Comparative electromagnetic articulographic study of English rhythm as produced by native and non-native speakers. Donna Erickson (1-11-1 Kamiasao Asao-ku, Kawasaki City, Kanagawa Prefecture, 215-8558 Japan), Atsuo Suemitsu (Japan Adv. Inst. of Sci. and Technol., Ishikawa, Japan), Yoshiho Shibuya (Kanazawa Medical Univ., Kanazawa, Japan), Sungbok Lee (USC, Los Angeles, CA), and Yu Tanaka (UCLA, Los Angeles, CA)

Previous work suggests that concurrent changes in jaw displacement and formant frequencies may manifest rhythm of spoken American English utterances [e.g., Erickson, in *Proceedings of the Speech Prosody 2010*, Chicago (May 2010), p. 1]. This paper compares jaw displacement and

corresponding formant frequencies of monosyllabic American English words produced on low vowels in four word phrases with varying positions of emphasis as spoken by three native speakers and three Japanese speakers of English. Preliminary findings suggest an alternating pattern of strong-weak jaw displacement along with corresponding formant changes for the native speakers. This pattern was not consistently seen for the non-native speakers and seemed to vary as a function of their skill-level in spoken English.

10:50

5aSCa9. Rhythmic conversion between speakers of different dialects. Jelena Krivokapic (Dept. of Linguist., Yale Univ., 370 Temple St., 302, New Haven, CT 06520-8366, jelena.krivokapic@yale.edu)

Speakers' convergence to each other's production of segments has been well established [Sancier and Fowler, J. Phon. **25**, 421–426 (1995); Nielsen, J. Phon. **39**, 132–142 (2011)]. Less is known about prosodic convergence, but it has been identified for stress and pitch accent [Krivokapic, JASA, **127**, 1851 (2010); Krivokapic, JASA **129**, 2658 (2011)]. In an acoustic experiment, rhythmic conversion between speakers of American and Indian English is examined. The two languages differ both in the stress patterns of individual lexical items and in their global rhythmic properties, with Indian English being syllable-timed and American English stress-timed. Changes in speakers' productions are examined using a synchronous speech task [Cummins, ARLO **3**, 7–11 (2002), Zvonik and Cummins, Proc. Eurospeech **2003**, 777–780 (2003)], where two speakers read sentences at the same time. Eight subjects (four dyads), each consisting of one Indian and one American speaker, read a short story that contained 12 words in which the two dialects differ in stress pattern and four sentences in which the global rhythmic properties are tested. The data of the dyad examined to date indicate convergence to a more stress-timed pattern for the Indian speaker.

11:05

5aSCa10. Analysis of rhythmic entrainment in speech production using real-time magnetic resonance imaging. Louis Goldstein, Michael I. Proctor, and Adam Lammert (Univ. of Southern California, Dept. of Linguist., Los Angeles, CA 90089)

It has been shown that subjects may recruit extra-linguistic articulators during the rhythmic production of speech (Tiede, 2010), that sympathetic head movement is associated with dysfluencies in running speech (Hadar, 1984), and that the magnitude of this activity positively correlates with both speaking rate and increased effort in accelerating speech production tasks (Hadar, 1991; Tiede, 2011). It is unclear at what level of linguistic organization this entrainment occurs. We examined patterns of articulation and head movement during an accelerated repetitive speech task involving multi-gestural segments, in order to gain more insights into the coordinative bases of this behavior. Real-time magnetic resonance imaging was used to examine the production of paired English words, contrasting along various articulations, e.g., "cop-top," "kid-kim," "flee-free," "muck-duck." Subjects' upper airways were imaged in the midsagittal plane while producing trochaic repetitions of each word pair in time to an accelerating metronome. Labial, tongue tip, tongue body and velic activity, as well as gross head movement, were tracked with direct image analysis (Lammert, 2010). Spontaneous head nodding was frequently observed in coordination with the prosodic foot. Nodding was observed to increase in amplitude with speech rate and during production of stimuli containing complex segments. [Work supported by NIH].

11:20

5aSCa11. Discriminating language and talker using non-linguistic measures of rhythm, spectral energy and f0. Kathy M. Carbonell (Dept. of Speech, Lang. & Hearing Sci., Univ. of Arizona, 1131 E 2nd St., Tucson, AZ 85721, kathyc@email.arizona.edu), Kaitlin Lansford, Rene Utianski, Julie Liss (Arizona state Univ., Tempe, AZ 85287-0102), and Andrew Lotto (Univ. of Arizona, Tucson, AZ 85721)

Recent studies have shown that rhythm metrics calculated from amplitude envelopes extracted from octave bands across the spectrum (the envelope modulation spectrum or EMS) can reliably discriminate between spoken Spanish and English even when produced by the same speakers [Carbonell *et al*., J. Acoust. Soc. Am. **129**, 2680]. Additionally, bilingual speakers could be discriminated fairly *seven females and five males* well on EMS variables even across sentences spoken in the different languages. In the current study, EMS, a general acoustic measure with no reference to phoneme/linguistic entities, was supplemented with measures of the mean and variance of spectral energy in each octave band as well as the mean and variance of fundamental frequency. Using stepwise discriminant analysis and the set of bilingual productions of Spanish and English, it was determined that language discrimination was excellent using both EMS and spectral measures, whereas spectral and f0 measures were most informative for speaker discrimination. The results demonstrate that this suite of easily-calculated acoustic measures provides abundant information about differences between languages as well as inherent differences in speakers that are independent of language spoken. [Work supported by NIH-NIDCD.]

11:35

5aSCa12. Temporal regularity in speech perception: Is regularity beneficial or deleterious?. Eveline Geiser (Dept. of Brain and Cognitive Sci., Massachusetts Institut of Technoloy, 43 Vassar St. 46-4033, Cambridge, MA 02139, egeiser@mit.edu), Stefanie Shattuck-Hufnagel (Res. Lab. of Electrons, MIT, Cambridge, MA 02139), and John D. E. Gabrieli (Dept. of Brain and Cognitive Sci., Massachusetts Institute of Technology, Cambridge, MA 02139)

Speech rhythm is of crucial importance for correct speech perception and language learning; for example, the specific rhythm is among the first things infants learn about their native language (Ramus, 2000). This study aimed to investigate the importance of speech rhythm in second language learning. German pseudosentences were presented to subjects in two conditions: spoken with a normal conversational speech rhythm versus with a speech rhythm containing a temporally regular beat. Nine native English speakers with 3.5 ± 1.6 years of German training repeated each sentence after hearing it once over headphones. Responses were transcribed using to the IPA and analyzed for the number of correct, false, and missing consonants, as well as for consonant intrusions. The over-all number of correct reproductions of consonants did not differ between the two experimental conditions. However, the experimental condition comprising a normal conversational speech rhythm resulted in significantly fewer consonant intrusions than the condition containing a beat. These results highlight the importance of speech rhythm in language perception/production and suggest, in particular, that second language learning may be facilitated by language-specific temporal grouping. [During the preparation of this work E. G. was supported by the Swiss National Science Foundation (SNF;PBZHP1-123304).]

11:50–12:10 Panel Discussion

Session 5aSCb

Speech Communication: Speech Perception by Native and Non-Native Speakers (Poster Session)

Grant McGuire, Chair
Dept. of Linguistics, Univ. of California, Santa Cruz, Santa Cruz, CA 95064

Contributed Papers

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

5aSCb1. Stimulus direction predictability dampens auditory neural responses to pitch-shifted voice feedback. Oleg Korzyukov, Lindsey Sattler, Roozbeh Behroozmand, and Charles Larson (Commun. Sci. and Disord., Northwestern Univ., 2240 Campus Dr., Evanston, IL 60208)

Models of neural mechanisms of voice control propose that voice auditory feedback is compared with an internal representation of the predicted voice output that is based on efference copies of motor commands. In addition, sensory memory from previous productions can also help the brain form predictions about incoming auditory feedback during vocal production. Previous studies have shown that the auditory evoked neural responses are maximally dampened in conditions where the incoming feedback closely matches the internal predictions. The present study aimed to determine if the predictability of stimulus direction influences the ERP responses to pitch shifts in voice auditory feedback. Subjects were tested with all upward stimuli, all downward stimuli, or randomized stimulus directions in separate blocks of trials. The N100 ERP response had a greater amplitude in conditions where the stimulus direction was randomized (unpredictable) compared to constant direction stimulus conditions (predictable), regardless of whether the stimuli were upwards or downwards. These findings suggest that auditory neural responses to predictable stimulus direction are dampened compared with unpredictable stimuli possibly because the predictable stimuli are suppressed by the internal predictions formed by the efference copy and/or sensory memory mechanisms.

5aSCb2. Duration of American English vowels and its effects on intelligibility for native and non-native speakers. Chia-Tsen Chen, Chang Liu, and Su-Hyun Jin (Dept. of Commun. Sci. and Disord., Univ. of Texas at Austin, Austin, TX 78712)

Durations of 12 American English vowels were measured for English-, Chinese-, and Korean-native speakers. Results showed that vowel duration was significantly affected by speaker groups. That is, vowels produced by English- and Korean-native speakers were significantly longer than vowels produced by Chinese-native speakers. The patterns of vowel duration as a function of vowel category were quite similar between native and non-native speakers. However, Mandarin Chinese vowels did not differ in duration across Chinese-native speakers, while durations of Korean vowels were significantly different across Korean-native speakers. These results suggest that non-native speakers were able to follow the duration pattern of native speakers in their vowel production. When vowel duration was equalized across vowel category with the central vowel nucleus remained, vowel intelligibility significantly dropped about 5% compared to vowels presented in the syllabic context for native and non-native speakers. The perceptual effects of vowel duration on vowel intelligibility appeared to be the same for native and non-native speakers, independent of speaker language background and the second language proficiency level.

5aSCb3. Perceptually relevant information in energy above 5 kHz for speech and singing. Brian B. Monson, A. Davi Vitela, Brad H. Story, and Andrew J. Lotto (Dept. of Speech, Lang., and Hearing Sci., Univ. of Arizona, 1131 E. 2nd St., Tucson, AZ 85721, bbmonson@email.arizona.edu)

Although high-frequency energy (HFE) is typically ignored in the speech perception and music literature, previous work has demonstrated that listeners are able to extract information from the structure of energy above 5 kHz. When exposed to HFE stimuli, listeners could accurately determine the mode of production (speech versus singing) and the gender of the speaker/singer. However, HFE in these previous experiments was presented in isolation at levels higher than what is typical of normal speech and singing. In the series of experiments presented here, the stimuli were presented at more appropriate levels. Additionally, speech-shaped masking noise from 0 to 5.6 kHz (set at the mean levels for an average speaker/singer) was used to determine if HFE is usable by listeners even in the presence of masking by the more intense energy below 5 kHz in typical speech/voice listening conditions. Results indicate that even when the HFE stimuli are presented at typical levels and in the presence of a masker, listeners are still able to discriminate production modes and gender. These results suggest that HFE contains information that normal-hearing listeners can potentially access during the perception of speech and singing. [Work supported by NIH-NIDCD.]

5aSCb4. Stability of compensatory behavior for real-time perturbations of vowel formants. Takashi Mitsuya (Dept. of Psych., Queen's Univ., 62 Arch St., Kingston, ON K7L3N6, Canada, takashi.mitsuya@queensu.ca), Ewen N. MacDonald (Dept. of Psych. & Elec. and Comput. Eng., Queen's Univ.), and Kevin G. Munhall (Dept. of Psych. & Otolaryngol., Queen's Univ.)

While we talk, not only do we listen to the speech sounds other people make but also we listen to our own voice in order to control the phonetic/phonological details of the sounds. One example of this relationship is demonstrated by real-time formant perturbation studies which show that talkers automatically change their formant production when auditory feedback does not match with the vowel they intend to produce. The reported results are consistent across studies, yet the variability between talkers is usually quite large, with some talkers showing a large magnitude compensation while others compensate modestly. To date, the degree to which talkers compensate has been assumed to be stable, but this has never been directly examined. The current study tested the stability of compensatory behavior for perturbed formant shifted feedback by repeatedly testing a group of talkers over the course of a few weeks to measure the variance in formant values and compensation across experimental sessions. The results will be discussed in terms of sensorimotor adjustment and speech production models.

5aSCb5. Overall pitch height as a cue for lexical tone perception. Jing Shen, Diana Deutsch (Dept. of Psych., Univ. of California San Diego, La Jolla, CA 92093, jshen@psy.ucsd.edu), and Jinghong Le (School of Psych., East China Normal Univ., Shanghai 200062, China)

Absolute pitch has been hypothesized to be involved in processing lexical tones in tone languages, which associate pitch information with verbal labels. Since possessors of absolute pitch in music utilize the overall pitch of a tone in making identification judgments, the hypothesis was tested that native speakers of Mandarin utilize the overall pitch of a lexical tone as a cue to retrieve its tone label. In a reaction time task, Mandarin syllables in all four tones were presented both in their original forms and also transposed to four different levels of pitch height; subjects listened to each token and judged whether or not its meaning corresponded to the original. It was found that although Mandarin syllables with transposed overall pitch heights were judged to be the same tones as the original tokens, subjects were significantly slower in making judgments for those syllables that were transposed to different levels of pitch height, compared to those that were presented at their original pitch heights. This effect was most extreme for tones 1 and 3. These findings suggest that overall pitch height serves as a cue for identifying lexical tones and further strengthens the link between absolute pitch and lexical tone perception.

5aSCb6. Perceptual importance of the voice source spectrum from H2 to 2 kHz. Jody Kreiman (Head and Neck Surgery, UCLA School of Medicine, 31-24 Rehab. Ctr., Los Angeles, CA 90095-1794, jkreiman@ucla.edu) and Marc Garellek (UCLA, Los Angeles, CA 90095)

Modeling the source spectrum requires understanding of the perceptual importance of different spectral-domain attributes of the voice source. Although the roles of H1-H2 and high-frequency harmonics in quality perception are somewhat understood, the extent of spectral detail that is perceptually significant in other frequency ranges is not known. This experiment examined the perceptual importance of amplitude difference between the second and fourth harmonics (H2-H4), which varies with linguistically significant changes in mode of phonation and pitch. To further determine if sensitivity depends on H1-H2 or on F0, we synthesized non-pathological male (F0=100 Hz) and female voices (F0 = 200 Hz) and manipulated the values of H2-H4 by 1 dB increments over a 20 dB range, for three different H1-H2 levels. Listeners heard pairs of stimuli in a 1-up, 2-down adaptive paradigm, and variations in just-noticeable-difference for H2-H4 as a function of F0 and H1-H2 were assessed with analysis of variance (ANOVA). Discussion will focus on the extent to which H2-H4 sensitivity covaries with F0, H1-H2, and the shape of the spectrum above H4. [Work supported by NSF and NIH.]

5aSCb7. Perceptual benefits of English coda-stop glottalization. Marc Garellek (Dept. of Linguist., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095, marcgarellek@ucla.edu)

Glottalized stops in codas (e.g., [t]) are commonly found in English, often when the coda-stop is unreleased. Glottalized phonation is known to increase the amplitudes of higher harmonics in the spectrum, possibly due to faster closing velocity of the vocal folds (Stevens, 1977). This suggests that glottalized phonation occurring before an unreleased stop may be useful to the listener for perceiving the formant transitions into the stop, because the higher frequencies will be amplified during glottalized phonation leading into the glottalized stop. To test this hypothesis, a phoneme monitoring task was conducted with native English speakers, who were asked to monitor for /t/ in onsets and codas of English words. The target stimuli were monosyllables ending in a coda-/t/, e.g., /bit/. All the stimuli were uttered by a phonetically trained speaker and varied according to whether the word was glottalized and/or ended in an unreleased stop. Preliminary results from the listeners' data suggest that their reaction time is faster for glottalized coda-/t/ (whether released or not), and that their detection accuracy improves when the coda-stop is both released and glottalized. Discussion will involve whether glottalization is a cue to formant transitions in general or to the presence of /t/ specifically.

5aSCb8. The role of vowel transitions in the perception of Mandarin sibilants. Yung-hsiang Shawn Chang (Dept. of Linguist., Univ. of Illinois, 707 S. Mathews, Ave., 4080E FLB, Urbana, IL 61801, yhchang@illinois.edu)

Vocalic formant transitions have been reported to play a role in identification of sibilants in English (e.g., Whalen, 1981), Shona (Bladon *et al*., 1987), and other languages. In distinguishing the Mandarin /s/-/ʂ/ contrast, while F2 at vowel onset is suggested to reflect the vocal tract configuration at the moment of fricative release into the vowel (Stevens *et al*., 2004), it has never been studied whether this acoustic cue contributes to phoneme perception. This study used a F2 continuum spliced between a fricative highly confusable between /s/ and /ʂ/ and the vowel /a/ for an identification task on 16 native listeners. The results did not show a categorical perception effect on /sa/ and /ʂa/ distinction along the F2 continuum. The second experiment used the naturally produced /sa/ and /ʂa/ whose vocalic transitions were cross-spliced to investigate if the listeners were affected by misleading formant transitions. Listeners' goodness ratings revealed no sensitivity to the acoustic manipulation. It was concluded that F2 is not a primary cue for the place distinction of Mandarin /s/ and /ʂ/. Our findings support Wagner *et al*.'s (2006) observation that the contribution of formant transitions to fricative perception is language-specific and depends on spectral similarity among target fricatives.

5aSCb9. Vowel nasality as a perceptual cue to pharyngeal consonants in moroccan arabic. Georgia, E. Zellou (Dept. of Linguist., 295 UCB, Hellems 290, Boulder, CO 80309-0295, georgia.weissman@colorado.edu)

Pharyngeal consonants have been shown cross-linguistically to be coarticulated with velopharyngeal port opening (Bladon and Al-Bamerni, 1982, Elgendy, 2001). We examined whether coarticulatory vowel nasality is a perceptual cue to an adjacent pharyngeal, and/or nasal, consonant in Moroccan Arabic (MA). Monosyllabic MA words spoken by a native speaker containing a pharyngeal, voiced /ʕ/ or voiceless /ħ/, or a nasal /n,m/ were cross-spliced with vowels from other contexts creating two conditions per word: "Oral" stimuli were pharyngeals cross-spliced with pharyngealized vowels (i.e., adjacent to /$d^{ʕ}$/) or nasals cross-spliced with oral vowels (i.e., adjacent to /d/). "Nasal" stimuli were cross-spliced with nasal vowels. The stimuli were presented to nine native MA speakers in a timed lexical repetition task. An RM ANOVA of repetition times showed a significant interaction between nasality, consonant, and directionality [$F(7,324) = 3.755$, p=0.046]. For anticipatory coarticulation, vowel nasality resulted in faster RTs, indicating facilitated perception for voiced /ʕ/ ($t(52) = 1.67$, p =.001) but not voiceless /ħ/. For carryover coarticulation, vowel nasality resulted in faster RTs for /ħ/ ($t(52)=2.05$,$p = .02$), but not /ʕ/. In both directions, vowel nasality resulted in faster RTs for nasals. These results indicate that MA listeners use vowel nasality as a perceptual cue to a word with a pharyngeal segment.

5aSCb10. Sharing the beginning versus sharing the end: Spoken word recognition in the visual world paradigm in Japanese. Hideko Teruya and Vsevolod Kapatsinski (Dept. of Linguist., 1290 Univ. of Oregon, Eugene, OR 97403)

Several studies have suggested that in English spoken word recognition, words sharing beginnings compete more than words sharing ends (i.e., the cohort competitor has an advantage relative to the rime competitor). We investigated the relative activation differences between the cohort competitor and the rime competitor in Japanese CVCV words using the visual-world paradigm (Allopenna *et al*., 1998). Results were analyzed using linear mixed effects models (Magnuson *et al*., 2008). Participants fixated referents of rime competitors more than those of cohort competitors when the cohort competitor shared only the first consonant with the target while the rime shared the rest of the word (VCV). When both competitors shared a syllable with the target, the rime competitor was fixated as much as the cohort competitor and the recognition of the target was delayed. The results suggest a syllabic or moraic basis to Japanese spoken word recognition, where lexical activation requires at least one mora of the word to match the signal. The weakness of cohort effects in Japanese challenge fully incremental models of word recognition. Whether the results would also hold for English is unclear as previous visual world spoken word recognition studies in English did not systematically manipulate degree of overlap.

5aSCb11. Low-level adjustments to gross-categorical mismatches in the perception of accented speech. Meghan Sumner (Dept. of Linguist., Stanford Univ., Stanford, CA 94305-2150, sumner@stanford.edu) and Samuel Tilsen (Cornell Univ., Ithaca, NY 14853-4701)

Variation in speech abounds. How do listeners understand a single word produced any number of ways as an instance of one particular word and not another? Sumner (2011) found that listeners perceive unaspirated voiceless stops [p] as a /p/ (rather than the native /b/) when exposed to more variation and when listeners are shifted down a VOT continuum in 5 ms steps. To understand this behavior, we compared order effects and noise in a phoneme categorization task in two conditions, both shifting listeners down a VOT continuum from a native /p/ (60 ms VOT) to a non-native /p/ (0 ms VOT). In one condition, listeners heard p-initial words with VOTS that made small steps down (5 ms) the continuum, with interspersed big jumps up (15–20 ms). In the other, listeners heard the same words with big jumps down and small steps up. Replicating Sumner (2011), listeners were successful in perceiving an unaspirated stop as voiceless. Interestingly, a difference between the conditions emerged, with greater boundary shifts made with big jumps down and small steps up. These results show (1) within-native-category fluctuations have no effect on categorization and (2) big differences in VOT serve as categorization anchors, while small steps drive perceptual adjustments.

5aSCb12. Investigating the effect of speech stimuli on temporal masking release. Evelyn Davies-Venn, Peggy Nelson, Yingjiu Nie, Adam Svec, and Katare Bhagyashree (Univ. of Minnesota, SLHS, 164 Pillsbury Dr., SE, Minneapolis, MN 55455)

The effect of masking release is still the source of numerous active investigations. However, differences in findings are sometimes noted among studies, especially related to potential gate frequency effects. The present study investigated the effect of masking release on listeners using a variety of speech materials . The main goal of this study was to investigate the effect of speech material on measures of masking release for listeners with normal hearing and hearing loss. The test stimuli were IEEE sentences and modified spondee words. To eliminate confounds of audibility and duration, the test stimuli were equated for duration and audibility. Listeners were tested across a wide range of audibility and gate frequencies. Performance and masking release results will be presented for these speech stimuli. [Work supported by NIDCD 008306.]

5aSCb13. Perception of Korean alveolar fricatives by American learners. Ahrong Lee and Hanyong Park (Dept. of Linguist., Univ. of Wisconsin-Milwaukee, 3243 N. Downer Ave. Milwaukee, WI 53211, ahrlee@uwm.edu)

This paper presents experimental data on how native English-speaking adult learners perceive Korean voiceless alveolar fricatives in a task soliciting sensitivity to prosodic context. College students in a second semester Korean class were asked to identify tokens of the two contrasting Korean fricatives, lax /s/ and tense /ss/, using letters of the Korean alphabet while rating their confidence on a scale from 1 (not confident at all) to 7 (very confident). Four native speakers of Korean ($M = 2$; $F = 2$) produced 24 tokens of the target sounds in construction with /a/ in three different prosodic contexts, viz., prevocalic word-initial, pre-emphatic intervocalic, and post-emphatic intervocalic positions. The results show a tendency for subjects to classify both Korean fricatives as lax /s/ under most conditions, except that tense /ss/ was often separately identified when preceded by an emphasized vowel. Phonetically, Korean /ss/ is of greater duration and intensity in a post-emphatic context than elsewhere, hence more distinct from /s/ in both Korean and English and so, apparently, more easily identifiable for beginning learners.

5aSCb14. English speakers' discrimination of Korean plain versus aspirated affricates. Jessica Maye and Jenna Silver Luque (Dept. of Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208)

The Korean three-way voicing distinction has been of much interest to phoneticians. Korean obstruents may be plain, tense, or aspirated, which differ in VOT, closure duration, and f0 of vowel onset (Cho *et al*., 2002; Kim, 1994). Previous research on Korean stops has shown that English speakers perceive all three categories as voiceless in initial position (Shin, 2001). The present study extends this to Korean affricates, which are partly cued by frication duration, an acoustic cue that is distinctive in English for differentiating fricatives from affricates (Repp *et al*., 1978). Twenty monolingual English speakers participated in the study. The stimuli were six tokens each of the plain and aspirated Korean affricates ([tSa] and [tSha]), naturally produced by a male native speaker of Seoul Korean. Participants completed an AX discrimination task, following five practice trials in which they received feedback on their answers. Results showed very poor discrimination that did not differ from chance (mean $d' = 0.892$; $p < 0.001$). These results indicate that despite the relevance of an additional acoustic cue that is distinctive in English (frication duration), the Korean contrast between plain versus aspirated affricates is difficult for English speakers to discriminate, paralleling previous research on Korean stops.

5aSCb15. The perception of Japanese pitch accent in "whispered" speech. Yukiko Sugiyama (Faculty of Sci. and Technol., Keio Univ., 4-1-1 Hiyoshi, Yokohama, Kohoku-ku, Yokohama 223-0061, Japan, yukiko_sugiyama@mac.com)

The perception of Japanese pitch accent was investigated using "whispered" speech in which the F0 was artificially removed. While the F0 is said to be the primary cue for pitch accent in Japanese, it is not certain whether acoustic correlates other than F0 exist. The results of previous production studies that examined vowel duration or devoicing as a correlate of pitch accent are not consistent. The present study attempts to find correlates of pitch accent from the other end, i.e., perception. A native speaker of Tokyo Japanese produced 14 disyllabic minimal pairs that differed only in the presence or absence of accent (e.g., /hana*/ "flower" when accented vs. /hana/ "nose" when unaccented) in a carrier sentence. The utterances were then edited by replacing the F0 by random noise, creating artificial "whispered" speech. Twenty-two native speakers of Tokyo Japanese identified the words they heard in two kinds of stimuli, the 14 minimal pairs as produced by the speaker and the whispered speech. The results suggest evidence of pitch accent in whispered speech. Implications of the results for the nature of Japanese pitch accent and the perception of accent will be discussed.

5aSCb16. Discrimination of Japanese affricate contrast by native speakers of Korean. Hiromi Onishi (Dept. of East Asian Studies, Univ. of Arizona, LSB 102, 1512 First St., Tucson, AZ 85721, honishi@email.arizona.edu)

It is known that Korean learners frequently experience difficulty producing the Japanese syllable /tsu/, which is often perceived as the syllable /tʃu/ by native Japanese speakers. The investigator administered an AXB discrimination test using these affricates (/tsu/-/tʃu/) in order to examine how well Korean native speakers can discriminate this contrast. These affricates were also tested in the shape of a whole syllable with the high-back vowel (/tsu/-/tʃu/). In addition, discrimination of the high-back vowel following the two affricates was also examined. The overall results suggest that native Korean speakers are capable of discriminating Japanese syllable /tsu/ from /tʃu/. However, the participants scored significantly lower for the vowel contrast than the syllable and consonant contrasts. Following the AXB test, an identification test was conducted using Hangul orthography in order to investigate the perceptual assimilation patterns. The results showed that the two syllables are assimilated to different syllables in Korean. The identification of the vowels showed that the /u/ in the contexts of /ts__/ was consistently assimilated to a single Korean category while some kind of vowel with preceding semivowel /j/ was selected most frequently for /u/ in the contexts of /tʃ__/. These results are discussed based on Best's perceptual assimilation model.

5aSCb17. Korean perception of the Japanese voiceless alveolar fricative /s/. Heather E. Simpson (Dept. of Linguist., Univ. of California Santa Barbara, Santa Barbara, CA 93106, hsimpson@umail.ucsb.edu)

Korean voiceless obstruents exhibit a relatively unique phonemic contrast based on laryngeal activity, a three-way tense, lax, and aspirated contrast for stops and affricates and a two-way tense and lax contrast for the alveolar fricative /s/. The relative primacy of various cues to the distinction, particularly for the affricate and fricative contrasts, remains controversial. Patterns in the adaptation of English loanwords have been used to infer the phonological basis for the mapping of English /s/ to Korean, however perception studies have shown that loanword adaptations do not fully account for the synchronic phonological mapping from English loans by Korean

native speakers (Schmidt 1996). Mapping from Japanese /s/ to Korean has been much less studied, but according to a loanword study by Ito *et al.* (2006), Japanese /s/ in word-initial and medial position is consistently mapped to Korean lax /s/, and geminate /s/ is consistently mapped to Korean tense /s*/. However, there appears to have been no comparable perception study to show synchronic mapping by Korean native speakers. The current study undertakes a preliminary evaluation of perceptual mapping of Japanese /s/ to Korean and finds that though the word initial Japanese /s/ conforms to loan-word predictions, word-medial /s/ does not.

5aSCb18. Non-native Japanese learners' perception of vowel length contrasts in Arabic and Japanese. Kimiko Tsukada (Dept. of Int. Studies, Macquarie Univ., NSW 2109, Australia, kimiko.tsukada@mq.edu.au)

This study examined the perception of short vs. long vowel contrasts in Arabic and Japanese by four groups of listeners differing in their linguistic backgrounds: native Arabic (NA), native Japanese (NJ), non-native Japanese (NNJ), and Australian English (OZ) speakers. Listeners' first languages differed in the extent to which vowel duration is used contrastively. In both Arabic and Japanese, vowel length is phonemic. English, on the other hand, utilizes vowel duration in a more limited way. Of interest was the discrimination accuracy of NNJ listeners who learned Japanese as a second language beyond childhood. As expected, the NA and NJ groups discriminated their native contrasts more accurately than all the other groups while the NNJ listeners showed a significant shift in their perceptual behavior and outperformed the OZ listeners who had no knowledge of Japanese in discriminating the Japanese vowel length contrasts. Furthermore, NNJ was the only group who showed a balanced pattern of discrimination accuracy in Arabic and Japanese. Taken together, the results obtained in this study suggest that NNJ learned to discriminate Japanese vowel length contrasts to some extent, but the learning did not carry over cross-linguistically to the processing of vowel length contrasts in an unknown language, Arabic.

5aSCb19. Perception of initial stops in tonal and non-tonal Korean. Hyunjung Lee and Allard Jongman (Dept. Linguist., Univ. of Kansas, 1541 Lilac Ln., Lawrence, KS 66045, hyunjung@ku.edu)

The current study investigated the perception of the three-way distinction among voiceless stops in non-tonal Seoul Korean and tonal Kyungsang Korean. We addressed whether listeners from these two dialects differ in the way they perceive the three stops. Twenty Korean listeners (nine from Seoul and eleven from Kyungsang) were tested with stimuli in which VOT and F0 were systematically manipulated. Analyses of the perceptual identification functions show that the results replicate the reported phonetic trading relationship between VOT and F0 in Seoul Korean and that this trading relationship is observed in Kyungsang as well. However, the trading relationship differs between the two dialects. Logistic regression analysis further shows that the two dialects use the perceptual cues differently. While both VOT and F0 are effective for Seoul Korean, only VOT is effective for the identification of all three stops for Kyungsang Korean. A similar pattern has been observed acoustically [Lee and Jongman, J. Acoust. Soc. Am. **127** (3), 2023–2023 (2010)]. The results will be discussed in terms of the close link between perception and production across the two different dialects.

5aSCb20. Lexical tone processing by monolingual and bilingual speakers of tone and non tone languages. D. Kyle Danielson (Dept. of Linguist., Univ. of Alberta, 4-32 Assiniboia Hall, Edmonton, AB T6G 2E7, Canada, dkd@ualberta.ca)

This experiment tested early Chinese–English bilinguals (age of arrival in Canada < 2 years) for lexical encoding processes of F0 contour (tone). Earlier experimentation (e.g., Mattock and Burnham, 2006) has indicated that processing of lexical tone differs between the two corresponding monolingual groups, demonstrating that English monolingual infants begin to disregard F0 contour differences in non word minimal pairs by the age of nine months, while Chinese monolingual infants do not exhibit this behavior. In the present study, three groups of participants were tested using a short-term memory encoding task (Dupoux *et al*., 2010) with non word minimal pairs differentiated only by F0 contour. Although the tone contours utilized in the experiment were non-native to all speakers, Chinese-dominant bilinguals (adult arrival in Canada) performed significantly better than English monolinguals in recalling long non word sequences differentiated only by these contours, while their performance in simple phoneme-differentiated sequences (e.g., [mu, fu]) was equal to that of the English speakers. However, the target group of early Chinese–English bilinguals produced scores corresponding to a bimodal distribution, with some speakers correlating to English monolinguals' performance and others corresponding to Chinese-dominant speakers' scores. Correspondence to one mode or another was analyzed using a series of sociolinguistic factors.

5aSCb21. Effect of vowel height and nasal coda on Mandarin tone perception by second language learners. Tzu-Yun Lai (Foriegn Lang. Dept., National Chiao Tung Univ., No. 513 Female 2 Domitory, 1001 University Rd., Hsinchu, Taiwan 300, Taiwan) and Yi-Wen Su (National Chiao Tung Univ., Taiwan 300, Taiwan)

The present study aims to investigate whether second language (L2) learners of Mandarin will be affected by vowel height sequences and post-vocalic nasals in perceiving Mandarin tones. A total of 101 stimuli from a female native speaker of Mandarin were recorded for constructing the main experiment. Eight Mandarin L2 learners and two native speakers were recruited. The result indicated that L2 learners had a higher error rate when perceiving H-L-H (high-low-high) and nasal coda. However, perceiving H-L (high-low) or L-H (low-high) in tone 2 and tone 3, L2 learners easily mix up these together.

5aSCb22. Effects of speaker variability and noise on identifying Mandarin fricatives by native and non-native listeners. Chao-Yang Lee, Yu Zhang, Ximing Li (Div. of Commun. Sci. & Disord., Ohio Univ., Athens, OH 45701), Liang Tao, and Z. S. Bond (Ohio Univ., Athens, OH 45701)

A fundamental issue in speech perception is how listeners with different characteristics deal with various sources of acoustic variability. This study examined how speaker variability and noise affected native and nonnative identification of the place contrast in voiceless Mandarin fricatives. Monosyllabic Mandarin words fa, sa, xia, and sha produced by three female and three male speakers were mixed with five levels of speech-shaped noise. The stimuli were presented either blocked by speaker or mixed across speakers to 40 native Mandarin listeners and 52 English-speaking listeners with various amounts of Mandarin experience. It was predicted that the mixed presentation and noise would affect nonnative identification to a greater extent because of the imperfect knowledge of the target language. The results showed that nonnative performance was compromised to a greater extent by noise but not by speaker variability. These results were compared to previous studies which compared native and nonnative perception of segmental and suprasegmental contrasts in other adverse conditions. It was concluded that not all sources of acoustic variability affect native and nonnative speech perception similarly.

5aSCb23. Listeners retune phoneme boundaries across languages. Eva Reinisch, Andrea Weber, and Holger Mitterer (Max Planck Inst. for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands, eva.reinisch@mpi.nl)

Listeners can flexibly retune category boundaries of their native language to adapt to non-canonically produced phonemes. This only occurs, however, if the pronunciation peculiarities can be attributed to stable and not transient speaker-specific characteristics. Listening to someone speaking a second language, listeners could attribute non-canonical pronunciations either to the speaker or to the fact that she is modifying her categories in the second language. We investigated whether, following exposure to Dutch-accented English, Dutch listeners show effects of category retuning during test where they hear the same speaker speaking her native language, Dutch. Exposure was a lexical-decision task where either word-final [f] or [s] was replaced by an ambiguous sound. At test listeners categorized minimal word pairs ending in sounds along an [f]-[s] continuum. Following exposure to English words, Dutch listeners showed boundary shifts of a similar magnitude as following exposure to the same phoneme variants in their native language. This suggests that production patterns in a second language are deemed a stable characteristic. A second experiment suggests that category retuning also occurs when listeners are exposed to and tested with a native speaker of their second language. Listeners thus retune phoneme boundaries across languages.

5aSCb24. Speech-in-noise perception for new sentence-recognition materials: Normative data for a diverse non-native listener population. Rajka Smiljanic (Linguist., Univ. of Texas at Austin, 1 University Station B5100 Austin, TX 78712-0198), Lauren Calandruccio, and Stacey Rimikis (The City Univ. of New York, NY 11357)

Even though the US population is becoming more diverse, there are no normalized sentence recognition tests for non-native English listeners. Recently, a large set of new test materials aimed at assessing speech

recognition and hearing abilities for native and non-native listeners were developed by Smiljanic and Calandruccio [J. Acoust. Soc. Am. **128**(4), 2486]. Five hundred unique sentences divided into 20 lists were created. The target keywords were chosen from the most frequent lexical items occurring in naturally elicited conversations with 100 non-native talkers with varied linguistic and cultural backgrounds. The lists were equalized for syntactic structure, syllable count, high-frequency phonemes, and word frequency. Perception results revealed equal mean performance across lists for native listeners. Currently, we are in the process of collecting normative data for these materials. Results of speech-in-noise tests for a diverse group of non-native English listeners with various native language backgrounds and levels of English experience will be presented. Regression analyses will be performed to investigate how the listeners' linguistic experience, competence, and proficiency relate to their listening performance on these new sentence-recognition materials. The results of this study add to our current understanding of non-native listener-related factors that affect speech recognition in adverse listening conditions.

5aSCb25. Phonemic perception under noise condition by early and late learners of English. Tetsuo Harada (Graduate School of Education, Waseda Univ., 1-6-1 Nishi Waseda, Shinjuku, Tokyo 169-8050, Japan, tharada@waseda.jp)

This study investigated long-term age effects of minimal input in an English as a foreign language (EFL) environment on the perception of English phonemes. Ten native speakers of American English, 10 university students who started studying English in Japan for a few hours a week between ages of three and eight (early learners), and 10 university students who began study in junior high school at the age of 12 or 13 (late learners) participated in a phonemic discrimination test. The selected target phonemes were tense vs. lax vowels ([i, ɪ], [u, ʊ]) and word-initial approximants ([l, r]). Three monosyllabic words were selected for each phoneme (e.g., *beat, bit*). Each pair of words was tested as six different tri-word trials (e.g., AAB, ABA) with and without the presence of a white noise. The participants were asked to choose the odd out in each trial. Results showed that although the early and late learners did not significantly differ across the phonemes in the discrimination test given without any noise, the former outperformed the latter under noise condition ($p < 0.05$). The findings support Lin *et al.*'s (2004) hypothesis that early learners are likely to establish more robust phonemic categories than late learners.

5aSCb26. Spanish learners' perception of American and Southern British English vowels. Paola Escudero (MARCS Auditory Labs., Bldg. 1, Univ. of Western Sydney, Bullecourt Ave., Milperra, NSW 2214, Australia) and Katerina Chladkova (Univ. of Amsterdam, 1012VT Amsterdam, The Netherlands)

Second-language (L2) studies show that learners differ in their vowel perception patterns, which can be attributed to their L2 proficiency or to different initial stages at the onset of L2 learning. The present study assessed the L2 perception of Southern British English (SBE) and American English (AE) vowels in 196 Spanish learners of English with different general comprehension proficiencies. Escudero and Chladkova (2010) showed that naive Spanish listeners perceived AE vowels differently than SBE vowels. The current results show that L2 learners' performance matches that of naive listeners. For instance, learners identified AE /ae/ correctly in only 3% of the time, as opposed to 32% for the same SBE vowel. This is because naive listeners assimilate AE /ae/ to Spanish /e/, which is the same native vowel they choose for AE /I/ and /E/, while SBE /ae/ is mostly assimilated to Spanish /a/. In addition, there was a significant effect of L2 proficiency on vowel identification accuracy and on the use of vowel duration. More proficient learners identified L2 vowels more accurately and used duration in a more native-like way than less proficient learners. Implications of our findings for models of L2 speech perception are discussed.

5aSCb27. Cross-language perception of Brazilian Portuguese vowels by Californian English speakers. Polina Vasiliev (Dept. of Spanish and Portuguese, UCLA, 5310 Rolfe Hall, Los Angeles, CA 90095, pvasiliev@ucla.edu) and Paola Escudero (Univ. of Western Sydney, Milperra, NSW 2214, Australia)

The present study investigates the perceptual assimilation and auditory categorization of Brazilian Portuguese vowels by monolingual speakers of Californian English. In the perceptual assimilation task, listeners classified 140 vowel tokens in terms of 10 native English vowel categories by choosing from written English words containing the vowels. The stimuli were isolated vowels extracted from nonce words produced by 10 male and 10 female native speakers of Brazilian Portuguese. In the auditory categorization task, the same listeners identified six Portuguese vowel contrasts which were presented in an XAB format, where X was a subset of the same natural vowel tokens from the previous task and A and B were synthesized prototypes of the Portuguese vowels. The results of the perceptual assimilation task demonstrate that Californian English listeners assimilate two out of six Portuguese vowel contrasts to more than two native vowel categories, resulting in many instances of multiple category assimilation (MCA). In the auditory categorization task, most of the contrasts with the lowest accuracy were those that showed MCA, which demonstrates that this pattern of assimilation may be the cause of Californian English listeners' perceptual difficulties with Portuguese vowels. Predictions for L2 acquisition are made.

5aSCb28. Perceptual confusability of French nasal vowels. Solene Inceoglu (Dept. of Second Lang. Studies, Michigan State Univ., UPLA 101, East Lansing, MI 48823, inceoglu@msu.edu)

The present study investigated native and non-native speakers' perception of the French nasal vowels [ɛ̃-õ-ɑ̃] and the oral vowel [o]. Thirty-four L1 American English intermediate learners of French and thirty-four French native speakers were asked to identify 60 one-syllable word items containing the four vowels [ɛ̃-õ-ɑ̃-o] in word-final position produced by a native speaker and randomly distributed across three conditions: audio-only, audio-visual and visual-only. Identification results revealed that overall performance was better in audio-visual and audio than in visual condition and that non-native speakers showed greater identification for [õ] than for [ɛ̃] and poorer identification for [ɑ̃]. A confusion matrix revealed that across conditions, the vowel [ɑ̃] resulted in the most misidentification and was often mistaken for [ɑ̃]. For both native and non-native speakers [õ] was sometimes confused with its oral counterpart [o]. The data also indicate different patterns of identification of the three nasal vowels in the three conditions. For instance the visual condition was particularly helpful for identifying [ɑ̃]. The obtained results will be discussed in relation to existing second language acquisition research studies.

5aSCb29. Timing of perceptual cues in Scots Gaelic. Natasha Warner, Andrew Carnie, Muriel Fisher, Jessamyn Schertz, Lionel Mathieu, Colin Gorrie, Michael Hammond, and Diana Archangeli (Dept. Ling., U. Arizona, Box 210028, Tucson, AZ 85721-0028, nwarner@u.arizona.edu)

Scots Gaelic, an endangered language, has several typologically unusual sound distinctions. Work on English or other languages cannot predict what perceptual cues Gaelic listeners might use to perceive these distinctions. The current work uses gating experiments, run in Scotland with 16 native listeners (monolingual in Gaelic until at least school age), to investigate timing of perceptual cues to the palatalization, preaspiration, and nasal frication contrasts. Results show that perceptual information about consonant palatalization is located primarily in the consonant itself, with weaker cues in the preceding vowel. For preaspiration, there is no perceptual information in the preceding vowel, but even half of the preaspiration provides a sufficient cue. The claimed nasal fricatives are particularly interesting, as true nasalized fricatives may be aerodynamically impossible, but nasalization could be realized on the preceding vowel, or without frication. The results show that listeners are only marginally able to hear this distinction, and that information does not increase through the signal: The little perceptual information present is already available during the preceding vowel. This confirms aerodynamic results showing some neutralization of this distinction and some nasalization during the preceding vowel. Overall, the results help us determine how information is conveyed using typologically unusual distinctions.

5aSCb30. Examining the role of variability in emotional tone of voice during online spoken word recognition. Maura L. Wilson and Conor T. McLennan (Dept. of Psych., Cleveland State Univ., 2300 Chester Ave., Cleveland, OH 44115, m.l.wilson90@csuohio.edu)

Previous studies demonstrate that listeners are faster to recognize words recently spoken by the same talker, relative to a different talker, when processing is relatively slow. The purpose of the present study was to examine intra-talker variability in emotional tone of voice on listeners' online

perception of spoken words. Based on previous work, we predicted reduced priming in the mismatch condition, when words are repeated in a different emotional tone of voice (e.g., sad and frightened), relative to the match condition, when words are repeated in the same emotional tone of voice (e.g., frightened)—when processing is relatively slow (experiment 2, hard lexical decision) and equivalent priming in the match and mismatch conditions when processing is relatively fast (experiment 1, easy lexical decision). The present study should provide a greater understanding of the role that variability in emotional tone of voice plays in listeners' online perception of spoken words.

5aSCb31. Perception of the proximal voice features for emotional speech. Sona Patel, Tanja Banzigar, and Klaus R. Scherer (NCCR Ctr. for Affective Sci., Univ. of Geneva, 7 Rue des Battoirs, CH-1205 Geneva, Switzerland, Sona.Patel@unige.ch)

The expression of emotions results in a number of changes to the acoustic signal. Considering that prosodic events may be encoded into the speech signal using multiple acoustic features, many acoustic variations may be redundant for listeners. An understanding of the acoustic changes that listeners can reliably use to differentiate emotional speech will help in the modeling emotions, for example, from a Brunswikian perspective. In the Brunswikian model of emotion communication, the acoustic features are represented as "distal cues" and the subjective perception of these changes are represented as "proximal percepts." Few attempts have been made to understand the proximal percepts of emotional speech in relation to the acoustic properties they represent. In this study, 19 listeners were asked to rate a set of 160 samples (2 sentences expressed in 8 emotions by 10 speakers) using 12 visual-analog scales (8 prosodic scales and 4 emotion dimensions). Results showed high reliability for all scales except roughness, articulation, and intonation. The emotions were especially differentiated on the dimensions of loudness, sharpness, and speech rate. Regression analysis was used to determine the underlying acoustic cues represented by these scales. The results of this analysis will be discussed.

5aSCb32. Perception of Canadian French word-final vowels by English-dominant and French-dominant bilinguals. Franzo Law II (Dept. of Psych., Univ. of Wisconsin-Madison, Waisman Ctr., 1500 Highland Ave, Madison, WI 53705, flaw@wisc.edu)

Self-identified English-dominant and French-dominant bilinguals from Montreal participated in a modified vowel identification task. Group differences in accuracy and speed for identifying experimental vowels /e, ɛ, o, u, y, ø/ were investigated relative to control vowels /i, a/, expected to be easiest and fastest to identify by both groups. Of interest was the performance on front-rounded /y-ø/ (non-phonemic in English) and /e-ɛ/ (phonologically contrastive in both languages, but // is disallowed word-finally in English). Both groups performed well overall in identifying experimental vowels although the French-dominant group was comparatively more accurate and faster. The English-dominant group was slower than the French-dominant group in identifying /y/ and /e/. Mouse cursor movements captured trial-by-trial revealed that both groups often moved the cursor toward the response button for /u/ before correctly identifying /y/. Results showed that English-dominant participants demonstrated less-automatic perception of most experimental vowels. However, performance speed and mouse patterns of the French-dominant group varied among native vowel categories, implying possible interactions between automaticity and auditory salience. Productions of /e-ɛ/ by participants were analyzed to explore the relative robustness in production of this contrast among participants. Correlations between task performance and measures of French proficiency are explored. [Work supported by NIH F31DC008075.]

5aSCb33. A perceptual study of monosyllabic Mandarin tones produced by three-year-old children growing up in Taiwan. Puisan Wong (Dept. of Otolaryngol., The Ohio State Univ., 915 Olentangy River Rd., Columbus, OH 43212)

Wong *et al*. (2005) reported that monolingual Mandarin-speaking three-year-old children in the U.S. have not produced adult-like Mandarin lexical tones in monosyllabic words. This study adopted the same methods to examine monosyllabic Mandarin tones produced by three-year-old children in Taiwan. Four hundred and thirty eight monosyllabic tones were collected from ten children and seven mothers in Taiwan. These productions and the 92 productions by the four U.S. mothers reported in Wong *et al*. (2005) were low-pass filtered to retain pitch information, but eliminate lexical information, and were presented to five Mandarin-speaking adults in Taiwan for tone categorization. U.S. mothers' tones were categorized with ceiling accuracies, similar to those judged by U.S. Mandarin-speaking adults reported in Wong *et al*. (2005). Taiwan mothers' tones were judged with lower accuracies, with the accuracy of tone 2 significantly lower than U.S. mothers'. Taiwan children's four tones were judged with 63%, 24%, 27%, and 74% accuracy, all significantly lower than US and Taiwan mothers except for tone 1. Children made more errors in tones 2 and 3 than tones 1 and 4. These findings were consistent with those in Wong *et al*. (2005), suggesting similar developmental patterns for children growing up in the U.S. and Taiwan.

5aSCb34. Context effects on reduced speech perception. Brian C. Ten Eyck (Dept. of Linguist., Univ. of Arizona, P.O. Box 210028, Tucson, AZ 85721, bteneyck@email.arizona.edu)

Acoustic information found in casual, conversational speech is often greatly impoverished and widely variable relative to more formal pronunciation. Previous work has highlighted the role of context in facilitating the comprehension of reduced speech at the individual word level; the experiment reported here measured the impact of varying available context on perceptual accuracy of multi-word speech reductions in American English. Forty-two subjects were exposed twice to sixteen 4–10 word reductions, the second time in one of four counterbalanced contextual conditions: (1) several seconds of audible preceding discourse; (2) an identical written transcript of the preceding discourse; (3) audible discourse of the pose-reduction clause ending; and (4) no additional context. Results were scored from subjects' typed responses based on the number of words correctly identified in each condition. By-subjects and by-items ANOVAs reveal that mean accuracy is correlated with the availability of contextual information. However, pairwise differences in mean accuracy across conditions indicate that discourse context in-and-of itself does not play a significant role in resolving multi-word speech reductions, while speaker-specific acoustic cues do. These results suggest that acoustic cues present in reduced multi-word utterances represent a preferred information source relative to syntactic or discourse context in interpreting massively reduced speech.

5aSCb35. Cross-language perception of Brazilian Portuguese vowels by Californian English listeners. Polina Vasiliev (Dept. of Spanish and Portugese, Univ. of California, Los Angeles, 5310 Rolfe Hall, Los Angeles, CA 90095, pvasiliev@ucla.edu) and Paola Escudero (Univ. of Western Sydney, Milperra, NSW 2214, Australia)

The present study investigates the perceptual assimilation and auditory categorization of Brazilian Portuguese vowels by monolingual speakers of Californian English. In the perceptual assimilation task, listeners classified 140 vowel tokens in terms of 10 native English vowel categories by choosing from written English words containing the vowels. The stimuli were isolated vowels extracted from nonce words produced by 10 male and 10

female native speakers of Brazilian Portuguese. In the auditory categorization task, the same listeners identified six Portuguese vowel contrasts which were presented in an XAB format, where X were a subset of the same natural vowel tokens from the previous task and A and B were synthesized prototypes of the Portuguese vowels. The results of the perceptual assimilation task demonstrate that Californian English listeners assimilate two out of six Portuguese vowel contrasts to more than two native vowel categories, resulting in many instances of multiple category assimilation (MCA). In the auditory categorization task, most of the contrasts with the lowest accuracy were those that showed MCA, which demonstrates that this pattern of assimilation may be the cause of Californian English listeners' perceptual difficulties with Portuguese vowels. Predictions for L2 acquisition are made.

5aSCb36. Training Korean listeners to perceive phonemic length contrasts in Japanese: Effects of speaking rate variation and contrast type. Mee Sonu (GITI, LASS Lab., Waseda Univ., Sophia Univ., 29-7 Bldg., 1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo 169-0051, Japan, sonumee@toki.waseda.jp), Keiichi Tajima (Hosei Univ., Tokyo 102-8160, Japan), Hiroaki Kato (NICT, Kyoto 619-0288, Japan), and Yoshinori Sagisaka (Waseda Univ., Shinjuku-ku, Tokyo 169-0051, Japan)

Japanese phonemic length contrasts are difficult to perceive for native-Korean listeners learning Japanese as a second language (L2). Aiming at an effective L2 training method for L2 learners, two experiments were conducted. Experiment 1 evaluated which acoustic cues Korean listeners rely on when categorizing the phonemic length contrast. Experiment 2 examined how differences in speaking rate variation (slow, normal, and fast) and contrast type (vowel contrast vs. consonant contrast) would affect the effectiveness of perceptual training, using a minimal-pair identification task with words embedded in carrier sentences. There were four training conditions comprising combinations of two contrast types (vowel or consonant length) and two speaking-rate variations (single rate or three different rates). Results show that L2 listeners exploit absolute segmental duration to identify phonemic length contrast rather than durational criteria that vary according to speaking rate. Moreover, the trained groups significantly improved in their overall accuracy after training. However, none of the training groups showed significant generalization to untrained contrast types. These results suggest that the effect of training is limited and does not generalize to untrained contrast types even when speaking rate variation is incorporated during training. [Work supported in part by the Grant-in-Aid for Scientific Research (B), JSPS.]

5aSCb37. Lexical and metrical cues for English word segmentation by Mandarin second language learners of English. Yu-Ting Luo and Chien-Tzu Liu (Dept. of Foreign Lang. and Lit., Natl. Chiao Tung Univ., 3/F, Humanities Bldg. 2, 1001 Ta-Hsueh Rd., Hsinchu 300, Taiwan)

This paper investigates the roles of lexical cues and metrical cues in word segmentation by Mandarin L2 learners of English with high and low English proficiency. A cross-modal form priming paradigm is adopted in the experiment. In the experiment, 33 participants (16 advanced learners, 17 beginning learners) heard a five-syllable phrase (e.g., anythingcorri, anything is the and corri is the ′), produced by a native American English speaker. After 100 ms, a three-syllable letter string (e.g., the corridor) was shown on the screen. The participants task was to decide whether a target was a real word. The contexts, primes, and targets have real word and non-word version and have either SW or WS (S: strong; W: weak) stress patterns. The results indicate that: (1) Lexical cues are the main strategy for advanced learners; (2) Beginning learners rely more on metrical cues; (3) When the stress pattern of contexts and primes are identical, it took participants less time to response. The results support previous studies that lexical and metrical are the cues listeners would rely on when doing segmentation. Furthermore, the results suggest that the strategies listeners rely on would change as they become more proficient in English.