

Session 2pSCc**Speech Communication: General Topics in Speech Communication II (Poster Session)**

Yukari Hirata, Cochair

Colgate University, Dept. of East Asian Lang. and Lit., 13 Oak Dr., Hamilton, NY 13346, USA

James Hillenbrand, Cochair

Western Michigan University, Dept. Speech Path. and Aud., 1903 W. Michigan Ave., Kalamazoo, MI 49008, USA

Natasha Warner, Cochair

University of Arizona, MPI Nijmegen, Dept. of Linguistics, Box 210028, Tucson, AZ 85721, USA

All posters will be on display from 3:40 p.m. to 5:20 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 3:40 p.m. to 4:30 p.m. and contributors of even-numbered papers will be at their posters from 4:30 p.m. to 5:20 p.m.

Contributed Papers

2pSCc1. A grouping approach to harmonic complexes. Dirkjan J. Krijnders (University of Groningen, Dept. Artificial Intelligence, P.O. Box 407, 9700 AK Groningen, Netherlands, j.d.krijnders@ai.rug.nl), Maria E. Niessen (University of Groningen, Dept. Artificial Intelligence, P.O. Box 407, 9700 AK Groningen, Netherlands, m.niessen@ai.rug.nl), Tjeerd C. Andringa (University of Groningen, Dept. Artificial Intelligence, P.O. Box 407, 9700 AK Groningen, Netherlands, t.andringa@ai.rug.nl)

Humans seem to perform sound-source separation for quasi-periodic sounds, such as speech, mostly on harmonicity cues. To model this function, most machine algorithms use a pitch-based approach to group the speech parts of the spectrum. In these methods the pitch is obtained either explicitly, in autocorrelation methods, or implicitly, as in harmonic sieves. If the estimation of pitch is wrong, the grouping will fail as well. In this paper we show a method that performs harmonic grouping without first calculating the pitch. Instead a pitch estimate is associated with each grouping hypothesis. Making the grouping independent of the pitch estimate makes it more robust in noisy settings. The algorithm obtains possible harmonics by tracking energy peaks in a cochleogram. Co-occurring harmonics are compared in terms of frequency difference. Grouping hypotheses are formed by combining harmonics with similar frequency differences. Consistency checks are performed on these hypotheses and hypotheses with compatible properties are combined into harmonic complexes. Every harmonic complex is evaluated on the number of the harmonics, the number of subsequent harmonics and the presence of a harmonic at the pitch position. By using the number of subsequent harmonics octave errors are prevented. Multiple concurrent harmonic complexes can be found as long as the spectral overlap is small.

2pSCc2. The role of generalization in learning perceptual contrasts. Meghan Sumner (Stanford University, Department of Linguistics, Margaret Jacks Hall, Stanford, CA 94305-2150, USA, sumner@stanford.edu)

Understanding spoken language is a difficult task. Variation caused by accents make this task even more daunting. This project examines the generalization of novel phonetic cues by listeners of non-native speech. Native speakers of English were trained on Korean-English pronunciations of minimal pairs with word-final contrasts (e.g., bed - bet). While native English speakers rely on vowel duration to cue a final voicing contrast, native Korean speakers of English do not produce a vowel duration contrast, but use a different distinguishing phonetic cue (e.g., voicing into closure duration). Listeners were trained on a subset of minimal pairs ending in final stops (bet - bed). Following successful training, listeners participated in an immediate

or long-term form priming task. Four conditions included: Trained Identity (bet - bet), Trained Non-identity (bed - bet), Novel Non-identity, Trained Rime (wet - wed), and Novel Non-identity, Novel Rime (beat - bead). Preliminary results show episodic treatment of words in the immediate priming task, but generalization to novel forms in the long-term priming task. These data suggest that early learning is mediated by episodic representations - learning proceeds episodically until sufficient information is available to support a more abstract generalization.

2pSCc3. Vocalizations of deaf infants before and after cochlear implantation. Mary K. Fagan (Indiana University School of Medicine, Riley Research Wing 044, 699 West Drive, Indianapolis, IN 46202, USA, mkfagan@indiana.edu)

Critical aspects of spoken language depend on perceiving and understanding speech sounds. For deaf infants, however, neither sound perception nor awareness of the sound-making consequences of their actions is available. The goal of this study was to uncover and document early, measurable effects of hearing loss on infant vocalization and changes in these behaviors following cochlear implantation. Participants were 8 deaf infants, 7-11 months old, and 8 infants with cochlear implants, 12-20 months old and 1-6 months post-implantation. Dependent variables include number of vocalizations, mean vocalization duration, and mean frequency before and after cochlear implantation and in relation to reduplicated babble onset and word production. Post-implant changes in vocal behaviors may have clinical implications for decisions regarding age of implantation and auditory habilitation as well as for understanding post-implant variability.

2pSCc4. Contrastive Focus in Taiwan Mandarin. Ho-Hsien Pan (Dept. of Foreign Languages and Literatures, National Chiao Tung University, 1001 TA Hsueh Road, 300 Hsinchu, Taiwan, hhpan@faculty.nctu.edu.tw), Hui-Ying Tsou (Dept. of Foreign Languages and Literatures, National Chiao Tung University, 1001 TA Hsueh Road, 300 Hsinchu, Taiwan, hysou@ms11.hinet.net), Wan-Ting Huang (Dept. of Foreign Languages and Literatures, National Chiao Tung University, 1001 TA Hsueh Road, 300 Hsinchu, Taiwan, caissybelle@yahoo.com.tw)

This study investigated the production and perception of given, new, and contrastive information in Taiwan Mandarin. Spontaneous dialogues elicited through a game were analyzed to access the acoustical parameters marking different information structures. In perceptual studies, two perceptual experiments, (1) a dialogue history experiment, of which listeners identified the questions preceding the sentences, and (2) a pair comparison experi-

ment, of which subject compared sentence with same wording but different information structures were used. The stimuli were either natural utterances form elicited spontaneous dialogues or resynthesized utterances with conflicting duration and f0 cues swapped from utterances of different information structures. Preliminary results of production experiments showed that duration elongation is a more consistent cue for marking narrow and contrastive foci than f0 range expansion. Listeners performed better in the pair comparison experiments than in the dialogue history experiment. For nature utterances, listeners were able to identify sentences with given information, given versus new information, and given versus contrastive information, but were not able to identify utterances with more than two contrastive foci. Listeners' performances declined when resynthesized utterances with conflicting duration and f0 cues were presented. When presented with conflicting cues, listeners relied more on durational than f0 cues.

2pSCc5. Comparison of Japanese expressive speech perception by Japanese and Taiwanese listeners. Chun-Fang Huang (Japan Advanced Institute of Science and Technology, 1-1, Asahidai, Nomi, 923-1292 Sendai, Japan, chuang@jaist.ac.jp), Donna Erickson (Showa University of Music, 808 Sekiguchi, Atsugi, 243-8521 Kanagawa, Japan, EricksonDonna2000@gmail.com), Masato Akagi (Japan Advanced Institute of Science and Technology, 1-1, Asahidai, Nomi, 923-1292 Sendai, Japan, akagi@jaist.ac.jp)

Language is an important tool in speech communication. Even without the understanding of one language, we can still judge the expressive content of a voice, such as happiness or sadness. However, sometimes misunderstanding of emotional communication occurs. It is not clear what the common/different features are that help or hinder people with different culture/native-languages background in making judgments about the expressivity of speech. In order to explore this question, we focus on Japanese and Taiwanese listeners perception of Japanese expressive speech utterances. We used the perceptual model proposed by [Huang and Akagi, *InterSpeech* 2005; 2007], which involves a concept called "semantic primitives"-- adjectives for describing voice perception. This concept simplifies and clarifies the discussion of common/different features in terms of acoustic cues and expressive speech perception categories. Using this model, we found some common/different features in the perception of expressive speech. Specifically, our results suggest that there may be primary and secondary semantic primitives associated with acoustic speech characteristics which are involved in the perception of expressive speech, and that people from different cultures/native-language background tend to use the same primary semantic primitives in perceiving expressive speech but different secondary ones.

2pSCc6. Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method. Hironori Takemoto (ATR Cognitive Information Science Laboratories, 2-2-2 Hikaridai, Seika-cho Soraku-gun, 619-0288 Kyoto, Japan, takemoto@atr.jp), Parham Mokhtari (ATR Cognitive Information Science Laboratories, 2-2-2 Hikaridai, Seika-cho Soraku-gun, 619-0288 Kyoto, Japan, parham@atr.jp), Tatsuya Kitamura (Konan University, Okamoto 8-9-1, Higashinada, 658-8501 Kobe, Japan, t-kitamu@konan-u.ac.jp)

An acoustic simulator based on the finite-difference time-domain (FDTD) method was evaluated by acoustic measurements on solid models of the vocal tract. Three-dimensional vocal tract (3D VT) shapes for a male subject during production of the five Japanese vowels were measured by magnetic resonance imaging. Transfer functions of the 3D VT shapes were computed by the acoustic simulator. The accuracy of the finite-difference algorithm was second-order in time and fourth-order in space. From the same 3D VT shapes, solid models were made of epoxide resin by a stereolithographic technique, and their acoustic transfer functions were measured using a time-stretched pulse signal. The calculated and measured spectra were compared up to 8 kHz. Although locations of major poles and zeros were common between calculated and measured spectra, the calculated spectra showed a larger number of small zeros, possibly caused by underestimation of VT acoustic losses in the simulation. However, the lower four formants were simulated remarkably accurately, with a mean absolute error

of only 2.2 % compared with the acoustic measurements. These facts indicated that the acoustic simulator can reasonably account for acoustic phenomena within the VT.

2pSCc7. Speech intelligibility determined with various tests presented against noise. Edward Ozimek (Umultowska 85, 61-614 Poznan, Poland, ozimaku@amu.edu.pl), Dariusz Kutzner (Umultowska 85, 61-614 Poznan, Poland, konsbol@wp.pl), Anna Warzybok (Umultowska 85, 61-614 Poznan, Poland, a_warzybok@poczta.onet.pl)

The paper describes three Polish tests for assessment of speech intelligibility under noisy conditions: the matrix test, the sentence test and the digit triplet test. The matrix test contains semantically unpredictable sentences of fixed syntactical structures. This test is very useful for long-term intelligibility measurements and is applicable for subjects with profound hearing loss or cochlear implants users. The sentence test comprises utterances taken from everyday speech and aims at clinical speech intelligibility measurements. The digit triplet test contains sequences of three digits that are spoken separately and is aimed at screening measurements, especially via telephone. The statistical and phonemic properties of the tests are analyzed and compared. The correlation between results obtained in the respective tests and the tests applicability are also discussed.

2pSCc8. Consonants, but not vowels, prime lexical decision following masked priming. Boris New (CNRS - Université Paris Descartes, 71 Avenue Edouard Vaillant, 92100 Boulogne-Billancourt, France, boris.new@univ-paris5.fr), Veronica Araujo (CNRS - Université Paris Descartes, 71 Avenue Edouard Vaillant, 92100 Boulogne-Billancourt, France, veronica.cruzaraujo@gmail.com), Nathalie Bour (CNRS - Université Paris Descartes, 71 Avenue Edouard Vaillant, 92100 Boulogne-Billancourt, France, psynat93@hotmail.com), Thierry Nazzi (CNRS - Université Paris Descartes, 45 rue des Saints-Pères, 75006 Paris, France, thierry.nazzi@univ-paris5.fr)

Recently, Nespor et al (2003) have argued that consonants play a more important role at the lexical level than vowels. This proposal has received direct support from studies showing that 16/20/30-month-old infants are better at processing specific consonantal than vocalic information while learning new words (Nazzi, 2005; Nazzi & New, 2007). This proposal is also supported, although indirectly, by studies on adults. Adults have been found to rely more on consonants than on vowels when asked to transform a nonword into a real word (Cutler et al., 2000). Moreover, adults presented with a continuous stream of syllables are able to segment it into "lexical units" by relying on consonantal but not vocalic information (Bonatti et al., 2005). In our study, we directly investigate this issue using a visual masked priming lexical decision task (an oral version being currently implemented). The test items are bisyllabic (CVCV or VCVC) and trisyllabic (CVCVCV or VCVCVC) words, which are presented following four different primes: identity (e.g., joli-joli), unrelated (vabu-joli), same consonants (jalu-joli) and same vowels (vobi-joli). Priming was found for the identity and same consonants conditions, not for the same vowels condition, supporting the privileged role of consonants at the lexical level.

2pSCc9. Frequency importance function of Mandarin Chinese speech. Jing Chen (Dept. of Machine Intelligence, Speech and Hearing Research Center, 2 Science Building, Peking Univ., 5 Yeheyuan Road, Haidian District, 100871 Beijing, China, chenj@cis.pku.edu.cn), Tianshu S. Qu (Dept. of Machine Intelligence, Speech and Hearing Research Center, 2 Science Building, Peking Univ., 5 Yeheyuan Road, Haidian District, 100871 Beijing, China, qutianshu@cis.pku.edu.cn), Xihong H. Wu (Dept. of Machine Intelligence, Speech and Hearing Research Center, 2 Science Building, Peking Univ., 5 Yeheyuan Road, Haidian District, 100871 Beijing, China, wxh@cis.pku.edu.cn), Qiang Huang (Dept. of Machine Intelligence, Speech and Hearing Research Center, 2 Science Building, Peking Univ., 5 Yeheyuan Road, Haidian District, 100871 Beijing, China, QIANG

.HUANG@SPREADTRUM.COM), Ying Huang (Dept. of Psychology, Peking Univ., 5 Yeheyuan Road, Haidian District, 100871 Beijing, China, innhuang@gmail.com), Liang Li (Dept. of Psychology, Peking Univ., 5 Yeheyuan Road, Haidian District, 100871 Beijing, China, liangli@pku.edu.cn), Huisheng S. Chi (Dept. of Machine Intelligence, Speech and Hearing Research Center, 2 Science Building, Peking Univ., 5 Yeheyuan Road, Haidian District, 100871 Beijing, China, chi@pku.edu.cn)

The speech intelligibility index (SII) theory objectively assesses speech intelligibility, and the frequency-importance function (FIF), which reflects the relative importance of various frequency bands to speech intelligibility for various languages, occupies the central part of the theory. However, the FIF has not been examined for tonal Chinese Mandarin speech. In this study, considering the characteristics of Mandarin speech, 50 phonemically-balanced one-syllable words were selected as the speech stimuli from the speech corpus "A Method for Calculating Articulation Index (GB/T 15508-1995)", which meets the National Standards of China, and the 1/3 octave FIF was measured and computed with the Fletcher's method. The results show that spectrum regions with frequencies 2000-4000 Hz are more important in Chinese speech than those in English speech. The results predicted by SII model with the new FIF fit the result of human psychophysical studies well, which indicate that the new FIF model is effective and appropriate for assessing the intelligibility of Mandarin speech even when a steady-state noise masker is co-presented. [Supported by the NSFC 60435010; 60535030; 60605016; 30670704]

2pSCc10. Arabic Diagnostic Rhyme Test using minimal pairs. Bachir Boudraa (USTHB Faculté d'électronique et d'informatique, BP 32, El-Alia, 16031 Alger, Algeria, b.boudraa@yahoo.fr), Malika Boudraa (USTHB Faculté d'électronique et d'informatique, BP 32, El-Alia, 16031 Alger, Algeria, mk.boudraa@yahoo.fr), Bernard Guerin (INPG, Avenue Félix Viallet, 38031 Grenoble, France, Bernard.Guerin@enserg.fr)

In this paper, we try to adapt to Arabic the diagnostic rhyme test (DRT) developed by Voiers for English and adapted to French by Peckels & Rossi. DRT is specific to each language and is essential to subjective assessments of coders and synthesizers. Six Arabic pertinent dimensions are used which are acuteness, compactness, tenseness, stridence, nasality and flatness. 72 monosyllabic meaningful pairs of words called minimal pairs have been developed. The apprehensibility of every attribute is tested in each of six vowel contexts. Every feature has been repeated six times and has been paired twice with every vowel.

2pSCc11. A fascinating personage in the history of phonetics: The phonetics and the politics of Elise Richter. Natasha Warner (University of Arizona, MPI Nijmegen, Box 210028, Dept. of Linguistics, Tucson, AZ 85721-0028, USA, nwarner@u.arizona.edu)

Elise Richter (1865-1943) was the first woman lecturer or professor in Austria or Germany, and she was also an experimental phonetician and even an experimental phonologist. Richter's career was delayed by the roles allowed for women at the time, and the end of her career, and her life, came about through the Nazis' policies toward Jews. Richter began her scientific work in Romance philology, but decided to seek explanations for historical sound change in phonetics. This led her to a productive career in experimental phonetics, and she even applied phonetic methods to study questions of phonological structure. Richter was a leader in exploring new methods and bringing together questions across widely differing fields. Although her specific experimental findings were superseded with time, Richter's refusal to be constrained by the limits of a field led her to ideas that have become current in phonetics just recently. In addition to her scientific career, Richter was politically active, both in working to expand women's opportunities and in general politics. This poster presents a new perspective on Richter's contributions to phonetics and on her view of feminism. Furthermore, it makes new observations of a parallel between Richter's political activities and her scientific work.

2pSCc12. The role of tongue articulation for /s/ and /z/ production in whispered speech. Hirohide Yoshioka (University of Tsukuba, 1-4-5 Shinmei, Urawa, 336-0023 Saitama, Japan, hirohide-y@mtc.biglobe.ne.jp)

Although the timing of the initiation and cessation of vocal fold vibrations is crucial to characterize the voiced and voiceless cognates, other cues, such as the duration of preceding vowels, the patterns of the formant transitions in the following vowels, and the period of stop closure, may also play important roles in performing these distinctions. The present study is to further investigate the role of tongue articulatory movements during voiced and unvoiced consonant production, specifically when the vocal folds do not vibrate during the production of fricative consonants, /s/ and /z/ in whispered speech. A normal Japanese speaker served as the subject. The palato-lingual contact patterns during the intervocalic consonant /s/ and /z/ in whispered speech were recoded using dynamic electro-palatography. The results show that the area of palato-lingual contact is clearly wider during /z/ production than /s/ production; the opening for fricative turbulent noise production is narrower and longer for /z/ than for /s/. In addition, the contact pattern is unstable for /s/ production.

2pSCc13. Production and perception of V1V2 described in terms of formant transition rates. René Carré (Laboratoire Dynamique du Langage, UMR 5596, CNRS, Université Lyon 2, 14 Avenue Marcelin Berthelot, 69363 Lyon cedex 07, France, recarre@wanadoo.fr)

Vowels can be produced with static articulatory configurations leading to stable formant frequencies (targets). Here, an algorithm computes area functions according to the criterion of minimal deformation leading to maximal acoustic variations. Within this evolutionary dynamics, the deformations of the tube are not performed to reach targets, unknown during the process, but to move in the acoustic space in order to increase acoustic contrast. The corresponding formant trajectories in the acoustic space can be described in terms of formant transition rates. For example, following this dynamic approach, to produce [ae] from [a], the transition rates of F1 and F2 are necessary and sufficient to represent [e] and at the very beginning of the transition and throughout the transition there is sufficient information to detect [e]. This assertion means that the transition duration is more or less constant. Studies of V1V2 production and perception characterized by their formant transition rates are presented. Such a representation leads to new interpretations of vowel reduction, coarticulation, perceptual overshoot, hyper and hypo speech, normalization.

2pSCc14. The effect of facial dynamics on infant perception of adult-directed speech in noise. Linda Polka (McGill University, School of Communication Sciences and Disorders, Beatty Hall, 1266 Pine Avenue West, Montreal, QC H3G 1A8, Canada, linda.polka@mcgill.ca), Robin Panneton (Virginia Polytechnic Instit. & State Univer., Williams Hall, Blacksburg, VA 24061-0436, USA, Panneton@vt.edu), Jessica Versele (Virginia Polytechnic Instit. & State Univer., Williams Hall, Blacksburg, VA 24061-0436, USA, Versele@vt.edu)

Recently Polka and Rvachew (in press) found that 6- to 8-month-old infants discriminated a native contrast ("bu" - "gu") when familiarized and tested with syllables presented in quiet, but not when familiarized with syllables mixed with a competing noise, even when testing was conducted in quiet. Because the competing noise (segments of cricket and bird song) did not create energetic masking, infant failure to encode the familiarization syllable was due to a disruption in attention to the speech stream. Importantly, in this study, infants watched a checkerboard while listening to syllables spoken in an adult-directed (AD) style. The current study investigated if the addition of a dynamic face facilitates infants' speech processing in the same task. Six-month-olds were habituated to repetitions of the native syllable "ba" mixed with the same noise but these infants saw and heard syllables produced in AD style. Following habituation, infants were presented familiar (repetitions of "ba") and novel test trials (repetitions of "ga"). Infants failed to show syllable discrimination as indexed by a reliable preference for

the novel test trials. These findings show that a dynamic face producing AD speech does not engage infant attention to speech enough to support syllable discrimination in noise.

2pSCc15. Voicing offsets and onsets in relation to intraoral pressure values in lingual obstruents of German. Laura L. Koenig (Haskins Labs. and Long Island Univ., Brooklyn, New York, NY 11201-8423, USA, koenig@haskins.yale.edu), Susanne Fuchs (ZAS, Schuetzenstr. 18, 10117 Berlin, Germany, fuchs@zas.gwz-berlin.de)

Phonation requires that tracheal pressure remain higher than intraoral pressure (Pio). In obstruent consonants, a major constriction in the upper vocal tract yields an increase in Pio, inhibiting phonation. The degree of Pio increase in consonants varies as a function of laryngeal and supraglottal apertures. Voiceless stops involve a rapid buildup and discharge of Pio, whereas fricatives involve more gradual changes in Pio. This work quantifies phonation offsets and onsets in German obstruents in relation to the Pio at these times. Pio signals were recorded via a pressure transducer affixed to the posterior end of an EPG palate while 9 speakers of standard German produced intervocalic voiceless consonant sequences (stops, fricatives, affricates, and clusters). Past theoretical work suggests that phonation offsets and onsets will show a hysteresis effect, with onsets requiring higher driving pressures than offsets. Of particular interest here is whether the extent of hysteresis differs among stops, fricatives, and obstruent sequences. Data on intraoral pressure change will also be compared with EPG data to explore how supraglottal constrictions affect Pio, and thus, phonation.

2pSCc16. Magnetic Resonance investigation of palatalized stop consonants and spirants in Russian. Galina Y. Kedrova (Moscow State Lomonosov University, GSP-2, Leninskije Gory, MGU, 1st Humanities, R. 983, Centre for New Technologies in Humanities, 119992 Moscow, Russian Federation, kedr@philol.msu.ru), Nikolay V. Anisimov (Moscow State Lomonosov University, GSP-2, Leninskije Gory, MGU, 1st Humanities, R. 983, Centre for New Technologies in Humanities, 119992 Moscow, Russian Federation, kedr@philol.msu.ru), Leonid M. Zaharov (Moscow State Lomonosov University, GSP-2, Leninskije Gory, MGU, 1st Humanities, R. 983, Centre for New Technologies in Humanities, 119992 Moscow, Russian Federation, kedr@philol.msu.ru), Yuriy A. Pirogov (Moscow State Lomonosov University, GSP-2, Leninskije Gory, MGU, 1st Humanities, R. 983, Centre for New Technologies in Humanities, 119992 Moscow, Russian Federation, kedr@philol.msu.ru)

The research in question is aimed at the experimental MRI-investigation of articulatory gestures corresponding to the palatalized Russian stop consonants [pʲ], [tʲ], [kʲ], [bʲ], [dʲ], [gʲ] and spirants [sʲ], [zʲ] [hʲ] countered with their non-palatalized counterparts [p], [t], [k], [b], [d], [g], [s], [z] [h] in Russian speech production. An experimental dataset was based upon 2D MR-images, audio- and video recordings taken from 4 native speakers of Russian (3 males and 1 female) producing VCCV sequences (Russian pseudo-words with the second vowel stressed) containing Russian consonants under investigation in the vocalic contexts [a]_[a] and [a]_[i]. All speaking subjects had standard pronunciation and were without any perceptible articulation disease. MRI investigation of the Russian consonantal phonemes was based upon admitted procedures and techniques though expanded with several new original methods elaborated by the Russian team of experimentalists. Experimental data was collected through several MRI sessions (done in a month and in a year with participation of the same speaking subjects), ensuring thereby credibility and robustness of the experimental results. A detailed analysis of the whole dataset of MR-images of palatalised Russian consonants countered with their non-palatalized counterparts revealed special articulatory pattern of palatalisation in Russian.

2pSCc17. Duration of Japanese singleton and geminate obstruents in two- to four-mora words. Yukari Hirata (Colgate University, 13 Oak Drive, Department of East Asian Languages and Literatures, Hamilton, NY 13346, USA, yhirata@mail.colgate.edu), Hiroaki Kato (ATR Cognitive

Information Science Laboratories, 2-2-2 Hikaridai, Seika-cho Soraku-gun, 619-0288 Kyoto, Japan, kato@atr.jp), Keiichi Tajima (Hosei University, 2-17-1 Fujimi, Chiyoda-ku, 102-8160 Tokyo, Japan, tajima@hosei.ac.jp)

Hirata and Whiton [J. Acoust. Soc. Am. 118, 1647-1660 (2005)] found an invariant durational structure for Japanese stop quantity distinction in two- and three-mora words across different speaking rates. The present study examined whether their finding extends to include fricative and affricate quantity distinctions and three- and four-mora words. Stimuli were stop, fricative, and affricate contrasts in four types of words, including (1) a long vowel (e.g., ka.so.o vs. ka.s.so.o), (2) a moraic nasal (e.g., ji.se.n vs. ji.s.se.n), (3) CV sequence (e.g., ho.so.ku vs. ho.s.so.ku), and (4) shorter words (e.g., i.shi vs. i.s.shi), spoken in isolation at three speaking rates by four native Japanese speakers. Duration of contrasting obstruents, words, and the interval between the onsets of the first and the second vowels (VOI) was measured. Results indicated durational patterns similar to those found previously for stop contrasts. In addition, duration of words, regardless of their segmental composition, reflected well the number of moras they contained. Finally, the ratio of the contrasting obstruent to the word and the ratio of the VOI to the mean mora duration were useful in classifying the singleton and geminate categories across rates. [Supported by Grant-in-Aid for Scientific Research, JSPS]

2pSCc18. Effects of hand gesture and lip movements on auditory learning of second language speech sounds. Spencer Kelly (Colgate University, 13 Oak Drive, Department of Psychology, Hamilton, NY 13346, USA, skelly@mail.colgate.edu), Yukari Hirata (Colgate University, 13 Oak Drive, Department of East Asian Languages and Literatures, Hamilton, NY 13346, USA, yhirata@mail.colgate.edu), Jen Simester (Colgate University, 13 Oak Drive, Department of Psychology, Hamilton, NY 13346, USA, jsimester@gmail.com), Jackie Burch (University of Rochester, 585 Elmwood Ave., Box 645, Rochester, NY 14642, USA, Jaclyn_Burch@mail.urmc.rochester.edu), Emily Cullings (Colgate University, 13 Oak Drive, Department of East Asian Languages and Literatures, Hamilton, NY 13346, USA, ecullings@mail.colgate.edu), Jason Demakakos (Colgate University, 13 Oak Drive, Department of Psychology, Hamilton, NY 13346, USA, jdemakakos@mail.colgate.edu)

Previous research found that auditory training helps native English speakers to perceive phonemic vowel length distinction in Japanese, but that their performance has never reached native levels (Hirata et al., 2007). Given that multimodal information, such as hand gesture and lip movements, influences semantic aspects of language processing and development (Kelly et al., 2002), we examined whether this multimodal information helps to improve native English speakers' ability to perceive Japanese vowel length distinction. Forty-five native English speakers participated in one of three types of training: (1) audio alone; (2) audio with hand gestures; and (3) audio with lip movements and hand gestures. Before and after training, participants were given phoneme perception tests that measured their ability to distinguish between short and long vowels in Japanese, e.g., /kato/ versus /kato:/. Our original prediction was that more modalities in training would result in greater learning. Although all three groups improved from pre- to post-test, there were no significant differences among the three training groups. Unlike the original prediction, hand gestures and lip movements did not seem to augment learning of difficult sound distinctions. We will discuss possible benefits and limitations of using multimodal information in second language speech learning.

2pSCc19. MRI-based 3-D vocal tract acoustic analysis of an American English lateral sound. Xinhui Zhou (Speech communication lab, Institute of systems research and department of electrical and computer engineering, University of Maryland, College Park, MD 20742, USA, zxinhui@umd.edu), Carol Y. Espy-Wilson (Speech communication lab, Institute of systems research and department of electrical and computer engineering, University of Maryland, College Park, MD 20742, USA, espy

@glue.umd.edu), Mark Tiede (Haskins lab and MIT Research lab of electronics, 300 George street suite 900, New Haven, CT 06511, USA, tiede@haskins.yale.edu), Suzanne Boyce (Department of communication sciences and disorders, University of Cincinnati, Mail location 0394, Cincinnati, OH 45267, USA, boycese@email.uc.edu)

The lateral sound /l/ in American English involves a complex articulatory configuration that includes one or two lateral channels along the sides of the tongue, and sometimes a lingual-alveolar contact. The acoustic characteristics of /l/ consist of a third formant (F3) that is close in frequency to the F3 of the adjacent vowel(s), but is often considerably weaker in amplitude. Generally, there is relatively weak energy in the F3-F5 region. The articulatory-acoustic relationship is not well understood. Previous studies did not consider the 3-D vocal tract model and the area function extraction was based on some assumed model. In this study, we constructed a 3-D vocal tract geometry based on magnetic resonance images from one subject during sustained production of syllabic dark /l/ (as in "pole"), and performed finite-element analysis. The harmonic analysis showed that the frequency response of the 3-D vocal tract is comparable in many ways to the spectrum of the acoustic signal, suggesting good agreement between the actual and reconstructed vocal tracts. A pole-zero pair is found in F3-F4 region, which may explain the weak energy level in this frequency region. Application of the 3-D wave propagation property to vocal tract area function extraction will also be discussed. [Research supported by NIH.]

2pSCc20. The acoustics of yer and non-yer vowels [e] and [o] in Slovak. Stefan Benus (Constantine the Philosopher University, Stefanikova 67, 94901 Nitra, Slovakia, sbenus@ukf.sk)

Yers of Slavic languages are vowels that alternate with zero and historically developed from high short lax vowels. In Slovak, both front and back yers were preserved and surface as [e] and [o] respectively. For example, the second [o] in *kotol* 'cauldron-Nom-Sg' is a yer because it disappears with adding a suffix vowel: *kotla* 'cauldron-Gen-Sg'. Compare with a non-yer [o] in *kostol* 'church-Nom-Sg', *kostola* 'church-Gen-Sg'. Traditional phonological accounts of this difference (e.g. Rubach 1993) assume that yer vowels are underlyingly different from non-yer vowels and merge with [e] and [o] through a phonological process. Therefore, these accounts predict that yer and non-yer vowels should be phonetically identical since they enter the phonetic component already merged as [e] or [o]. The results of our acoustic experiments show that yer vowels are phonetically different from non-yer vowels. The most salient differences were observed in the first formant and duration: yers have lower F1 than non-yers, and for some subjects they are also shorter. This finding supports the view that the phonetic component has access to deep phonological alternations and that phonetics-phonology is a single cognitive system in which the components have different granularities and interact bidirectionally.

2pSCc21. Behavioral and physiological correlates of language preference. Megha Sundara (UCLA Department of Linguistics, 3125 Campbell Hall, Los Angeles, CA 90095-1543, USA, megha.sundara@humnet.ucla.edu), Christine Kitamura (MARCS Auditory Laboratories, Bankstown Campus, Building 1, University of Western Sydney, Locked Bag 1797, 1797 Penrith South DC NSW, Australia, c.kitamura@uws.edu.au), Thierry Nazzi (CNRS - Université Paris Descartes, 45 rue des Saints-Pères, 75006 Paris, France, thierry.nazzi@univ-paris5.fr)

With increasing experience listening to language, infants' sensitivity to the patterns of their native language becomes more specialized. In monolingual, hearing infants, this tuning-in into the native language has been demonstrated in an elegant experiment. Hayashi et al. (2001) tested Japanese-learning infants between 4 and 14 months on their preference for listening to Japanese when compared to English. Not surprisingly, Japanese-learning infants preferred to listen to Japanese over English. Critically, the magnitude of this native language preference increased linearly from 4- to 14-months. In this study, we replicated and extended Hayashi et al.'s results. We tested 65 Australian-English learning infants aged 3 - 16 months on their preference for Australian-English compared to Japanese sentences. Our results confirm the findings of Hayashi et al.: Australian-English hearing infants' prefer their native language, and this preference increases linearly with age.

We also tested preference when the same infants heard Australian-English and German sentences. English and German are rhythmically similar, while differing in segmental characteristics. Results indicate that similar to early language discrimination, rhythmic distance affects the emergence of language preference. We are currently analysing concomitant heart rate measured from these infants when they were listening to native and non-native sentences.

2pSCc22. Improvement of speech recognition thresholds by spectral modulation enhancement. Chang Liu (University of Texas, Dept. of Comm. Sci. & Dis., 1 University Station A1100, Austin, TX 78712, USA, changliu@mail.utexas.edu), David A. Eddins (University of Rochester, Department of Otolaryngology, Rochester, NY 14618, USA, David_Eddins@URMC.Rochester.edu)

Our previous studies have shown that vowel identification in noise was significantly improved by spectral enhancement resulting from modifications in the spectral modulation domain. The present study investigates whether spectral enhancement results in improved speech recognition in background competition. Speech recognition thresholds (SRT) were measured in a long-term speech shaped (LTSS) noise and in multi-talker babble using NU-6 words with and without spectral enhancement for young normal-hearing listeners. Results indicated that SRTs were significantly reduced, by 2.2 dB, in LTSS noise and in babble when speech sounds were spectrally enhanced between 2.0 and 2.5 cycles/octave, while no significant changes in SRTs occurred for enhancement between 1.5 and 2 cycles/octave or between 1.5 and 2.5 cycles/octave. Spectral modulation transfer functions were also measured for each listener. Linear regression analyses showed that SRT improvement was moderately correlated with modulation detection thresholds at a spectral modulation frequency of 0.5 cycles/octave. Thus, it appears that better modulation detection at low spectral modulation frequencies is associated with greater improvement in word recognition in background competition.

2pSCc23. Perception of sinewave vowels. James M. Hillenbrand (Western Michigan University, Dept of Speech Path & Aud., 1903 W. Michigan Ave., Kalamazoo, MI 49008, USA, james.hillenbrand@wmich.edu), Michael J. Clark (Western Michigan University, Dept of Speech Path & Aud., 1903 W. Michigan Ave., Kalamazoo, MI 49008, USA, michael.clark@wmich.edu)

There is a significant body of research examining the intelligibility of sinusoidal replicas of naturally spoken utterances. Discussion has followed about what the sinewave speech phenomenon might imply about the auditory and perceptual mechanisms that are involved in phonetic recognition. However, this work has typically been conducted using meaningful and syntactically well formed sentences, making it unclear what the relative contributions are of higher level, top-down processes as opposed to lower level auditory and phonetic mechanisms. The simple purpose of this study was to measure vowel intelligibility using sinusoidal replicas of naturally spoken /hVd/ syllables. The sinusoidal signals were modeled after 300 utterances selected from a database of 1,668 syllables spoken by men, women, and children [Hillenbrand, Getty, Clark, and Wheeler, *J. Acoust. Soc. Am.*, 97, 3099-3111 (1995)]. Listeners consisted of 74 students enrolled in an introductory phonetics course. Recognition rates for the sinusoidal vowels averaged 55%, with considerable variability across listeners. This figure is substantially lower than the ~95% intelligibility (measured in several earlier studies) of the naturally spoken signals upon which the sinusoidal replicas were modeled. Attempts to improve performance using three different training methods met with modest success, with post-training recognition rates rising by ~5-12 percentage points, depending on the training method.

2pSCc24. Complexity of acoustic-production-based models of speech perception. Geoffrey Stewart Morrison (Australian National University, School of Language Studies, Building 110, ACT 0200 Canberra, Australia, geoff.morrison@anu.edu.au)

Discriminant analysis models trained on acoustic vowel production data have been found to have significant correlation with listeners' perception. Two regularised discriminant analysis models were trained on monolingual speakers' vowels. One model was trained on North Central Peninsular Span-

ish vowel tokens, and the other on Western Canadian English vowel tokens. For each language the model which resulted in the lowest cross-validated classification error rate was close to the least complex model possible, i.e., close to linear discriminant analysis using the variances of each acoustic variable but not the covariances between variables. In order to make cross-language vowel perception predictions the Spanish model was used to classify English vowel tokens and the English model was used to classify Spanish vowel tokens. Results suggest that monolingual North Central Peninsular Spanish listeners would assimilate most tokens of Western Canadian English /i/ and /I/ to Spanish /i/ and /e/ respectively, and thus for this combination of dialects, Spanish-speaking learners of English would not be expected to have difficulty with the English /i/-/I/ contrast.

2pSCc25. Letter sound and letter name recognition for automated literacy assessment of young children. Shrikanth Narayanan (University of Southern California, 3740 McClintock Ave, EEB430, Los Angeles, CA 90089-2561, USA, shri@sipi.usc.edu), A Kazemzadeh (University of Southern California, 3740 McClintock Ave, EEB430, Los Angeles, CA 90089-2561, USA, kazemzad@usc.edu), Matthew Black (University of Southern California, 3740 McClintock Ave, EEB430, Los Angeles, CA 90089-2561, USA, Matthew.Black@usc.edu), Joseph Tepperman (University of Southern California, 3740 McClintock Ave, EEB430, Los Angeles, CA 90089-2561, USA, tepperma@usc.edu), Sungbok Lee (University of Southern California, 3740 McClintock Ave, EEB430, Los Angeles, CA 90089-2561, USA, sungbokl@usc.edu), Abeer Alwan (University of California, 405 Hilgard Ave, Los Angeles, CA 90095, USA, alwan@ee.ucla.edu)

Evaluations of letter naming and letter sounding are commonly used to measure a young child's growing reading ability, since performance in them is well-correlated with future reading development. Assessing a child's oral reading skills requires teachers, as well as technologies that attempt to automate such assessment, to form an item-level accept/reject decision based on speech cues and prior knowledge of the child's literacy level and linguistic background. With data collected from 171 K-2 children, both learners and native speakers of American English, we designed and evaluated an automated letter naming assessment method using a simple word-loop HMM decoding for the word-level letter names. The automated accept/reject evaluation performance, 81.9%, approached the agreement of human raters, 83.2% (0.62 kappa). However, the task where children must produce the sound that the letter represents was more difficult: English orthography allows one-to-many letter-to-sound mapping, teachers showed less agreement in their assessment (80.9%, 0.55 kappa), and the brief durations of some of the letter sounds made it difficult to distinguish them from each other and from background noises. Phone-level HMM based evaluation accuracy was 58.2%. Preprocessing the recordings into speech, silence, and noise improved these results, especially for plosive sounds. [Supported by NSF]

2pSCc26. Neighbors as competitors: Phonological analysis of spoken word recognition errors. Adam Buchwald (NYU, 665 Broadway, Suite 910, New York, NY 10012, USA, buchwald@nyu.edu), Robert A. Felty (Indiana University - Speech Research Laboratory, 1101 E 10th St., Dept. of Psychology, Bloomington, IN 47401, USA, robfelty@indiana.edu), David B. Pisoni (Indiana University - Speech Research Laboratory, 1101 E 10th St., Dept. of Psychology, Bloomington, IN 47401, USA, pisoni@indiana.edu)

Theories of speech perception and spoken word recognition widely assume that phonetically similar words compete for selection during lexical access. These competitors - called lexical neighbors - are classically defined as differing by a single segment, with little attention paid to other levels of phonological structure such as featural content or syllabic structure. This paper reports on analyses of over 15,000 word recognition errors made on a representative sample of the lexicon (N=1428) that includes the same variety of syllable structures, lengths, and lexical frequencies that exists in the English lexicon. The analyses reveal that competitors share segmental information as has been previously suggested, but that the classic definition of neighbor cannot capture the level at which featural and syllabic properties of the target constrain the error responses. We will discuss several analyses on

these issues and discuss the implications for our understanding of the organization of the lexicon. [Work supported by NIH-NIDCD R01 00111 and T32 00012]

2pSCc27. The prosody of second position clitic placement and focus in Croatian. Kristine Yu (Dept. of Linguistics, University of California, 3125 Campbell Hall, Los Angeles, CA 90095, USA, krisyu@humnet.ucla.edu)

Since Browne 1974, the placement of second position clitics in Bosnian/Croatian/Serbian has inspired debate about interaction at the syntax-phonology interface. The placement of these clitics can alternate quite freely: either after the first phonological word or after the first syntactic constituent. While it's generally agreed that prosodic phonology, in addition to morphosyntax, plays a role in clitic placement in Bosnian/Croatian/Serbian, the prosodic patterns for the different placements have not been studied acoustically. In addition, it has been suggested that clitic placement and pragmatic focus may interact (Bošković 2001), but this has not been systematically studied. We recorded adult Zagreb Croatian speakers producing subject noun phrases with initially stressed trisyllabic adjectives and nouns in transitive sentences. We varied clitic placement (after the first word or first constituent) and focus domain (broad focus, and narrow focus on adjective, noun, both the adjective and the noun, or the entire noun phrase) and controlled for dialect and pitch accents. Preliminary results from three speakers indicate that gross differences in prosodic patterns occur only as a function of focus domain but not clitic placement. Further work will confirm if these patterns hold for more speakers and if clitic placement is correlated with finer-grained prosodic differences.

2pSCc28. English /r/-/l/ pronunciation training for Japanese speakers. Kota Hattori (University College London, Department of Phonetics and Linguistics, 4, Stephenson Way, NW1 2HE London, UK, k.hattori@ucl.ac.uk), Paul Iverson (University College London, Department of Phonetics and Linguistics, 4, Stephenson Way, NW1 2HE London, UK, p.iversen@ucl.ac.uk)

Previous studies have demonstrated that perceptual training improves both perception and production by adult second-language (L2) learners. The present study examined whether production training likewise improves both perception and production. Japanese speakers underwent ten sessions of production training for English /r/ and /l/. The training combined three methods: (1) subjects received audio-visual examples and pronunciation instructions for /r/ and /l/; (2) they produced mono syllables (i.e., /ra/ and /la/) and minimal-pair words, with online feedback about their formant frequencies using a real-time spectrogram as well as pronunciation feedback from an instructor; and (3) they made recordings of minimal-pair words and compared them to "enhanced" versions of these recordings (i.e., signal-processed versions with the correct F3 frequencies and transition durations). All participants completed a battery of pre/post training tests (i.e., English /r/-/l/ identification, discrimination, perceptual mapping of best exemplars, and production). The preliminary results indicate that this training approach improves production; the results will be further discussed in terms of its effect on perception and the underlying representation of these categories.

2pSCc29. Production of American English [eI] and [e] by Mandarin speakers: mono- vs. disyllabic words. Chung-Lin Yang (Dept. of Linguistics, Indiana Univ.-Bloomington, Memorial Hall 322, Bloomington, IN 47405, USA, cy1@indiana.edu)

Mandarin productions of English tense vs. lax vowels are difficult for English speakers to differentiate (Chen et al. 2006). In this study the production of American English [eI] and [e] by Mandarin and American participants was investigated. The target vowels were embedded in monosyllabic and disyllabic words with a stop-V-voiceless stop context in carrier sentences of variable length. There were two formant measurements for each vowel (20% after the onset and before the end of the vowel). In the monosyllabic condition, Mandarin speakers' performance was similar to American speakers in duration and formant values except that [eI] started too low. However, in the disyllabic condition, Mandarin speakers' formants for [eI] and [e] were overlapped and there was no significant durational distinction between [eI] and [e]. One possible account for this difference is that

when producing a disyllabic word, Mandarin speakers tend to shorten the vowel duration of the first syllable to compensate for the production of the following syllable. The difference in syllable conditions may reflect a tendency that Chinese regularizes word durations regardless of the number of syllables. Reference: Chen, Y. et al. (2006). Vowel production by Mandarin speakers of English. *Clinical linguistics and phonetics*, 15, 427-440

2pSCc30. A numerical analysis of fluctuations in pressure wave within the larynx using two-dimensional asymmetrical vocal folds model.

Hideyuki Nomura (Dept. of Electronic Eng., Univ. of Electro-Communications, 1-5-1, Chofugaoka, 182-8585 Chofu-shi, Japan, nomu@ee.uec.ac.jp), Tomoo Kamakura (Dept. of Electronic Eng., Univ. of Electro-Communications, 1-5-1, Chofugaoka, 182-8585 Chofu-shi, Japan, kamakura@ee.uec.ac.jp), Tetsuo Funada (Div. of Electronic Eng. and Computer Sci., Kanazawa Univ., Kakuma-machi, 920-1192 Kanazawa-shi, Japan, funada@t.kanazawa-u.ac.jp)

Numerical simulations of pathological voice production and estimations of pressure wave fluctuations are performed based on a two-dimensional asymmetrical vocal folds (VFs) model. The asymmetrical VFs model takes into account of geometrical asymmetries (the thickness, effective depth of vibration region, and lateral rest position) and mechanical asymmetries (the Young's modulus, density, and viscosity of VF tissues). Simulation results based on the asymmetrical VFs model show that the left and right VFs vibrate with a phase difference. Obtained pressure waves within the larynx and vocal tract indicate fluctuations of fundamental frequency, amplitude, and waveform. In order to quantitatively evaluate the fluctuations, the coefficient of variation of the fundamental frequency, the coefficient of variation of the amplitude, and the harmonic-to-noise ratio are estimated. With increasing the VF asymmetries, especially on the effective depth and the density of VF elements, remarkable fluctuations are observed not only at the glottis but also in the regions downstream of the glottis and faraway from the glottis. This suggests that the estimation of fluctuation in pressure wave is useful for the diagnosis of pathological VFs.

2pSCc31. Differential effects of the phonemes on identification of previously unknown speakers.

Kanae Amino (Dept. of Electrical and Electronics Engineering, Sophia University, 7-1 Kiyoi-cho, Chiyoda-ku, 102-8554 Tokyo, Japan, amino-k@sophia.ac.jp), Takayuki Arai (Dept. of Electrical and Electronics Engineering, Sophia University, 7-1 Kiyoi-cho, Chiyoda-ku, 102-8554 Tokyo, Japan, arai@sophia.ac.jp)

In perceptual speaker identification, it is known that the identification accuracy depends on the contents of the stimuli presented to the listeners. In our previous experiments, we found that the stimuli containing a nasal sound are effective for human speaker identification, and that coronal sounds are more effective than labial ones. This tendency was observed in the identifications of both familiar and previously unknown speakers. In this present study, we investigated the effects of the speech contents again, by focusing on the vowels in CV monosyllabic stimuli. Through the experiment we obtained several findings: 1) stimuli containing a nasal gained significantly higher accuracy compared to stimuli with only oral sounds; 2) coronal sounds were more effective than labial sounds; 3) palatalisation caused an improvement in performance; and 4) back vowels were more effective than front vowels significantly. These results can be explained by speaker-specific morphologies of the nasal cavity and the paranasal sinuses, and also of the pharyngeal cavity. We will also show analyses and discussions on the acoustical properties of the stimuli and the performance differences among the subjects. [Work supported by Sophia University Open Research Centre from MEXT.]

2pSCc32. Coarticulation in CV sequences: Locus Equation data.

Simone Graetzer (University of Melbourne, Royal Parade, Parkville Victoria, 3010 Melbourne, Australia, n.graetzer@pgrad.unimelb.edu.au)

Background: Speech sounds are known to be influenced by the context in which they occur, and to be produced in an overlapping way. These processes are referred to by the term "coarticulation". Objective: The aim of the study was to compare the magnitude of anticipatory coarticulation of oral voiceless plosives by following vowels in four Australian languages, Arrernte, Burarra, Gupapuyngu, and Warlpiri. Methods: The corpus consists of

approximately six hundred isolated real words spoken by eight female adult speakers of the four languages. Locus equations are calculated for intervocalic consonants /p,c,t,tr,k/ (where "/tr/" represents a voiceless oral retroflex plosive) and all following vowels in consonant-vowel sequences. Measurement points are 0.1 and 0.5 into the vowel. The relationship between these two measurement points is linear and is expressed as a locus equation. The slope of the fitted line associated with the equation is a measure of the magnitude of coarticulation of the consonant by the following vowels. Results: There was a general tendency towards relatively highly coarticulated bilabial and velar plosives, weakly coarticulated retroflex plosives, and moderately coarticulated alveolar plosives. Results were variable for palatal plosives. These results are for the most part in accordance with findings for other Australian languages.

2pSCc33. Lexical analysis of spoken word recognition errors.

Robert A. Felty (Indiana University - Speech Research Laboratory, 1101 E 10th St., Dept. of Psychology, Bloomington, IN 47401, USA, robfelty@indiana.edu), Adam Buchwald (NYU, 665 Broadway, Suite 910, New York, NY 10012, USA, buchwald@nyu.edu), David B. Pisoni (Indiana University - Speech Research Laboratory, 1101 E 10th St., Dept. of Psychology, Bloomington, IN 47401, USA, pisoni@indiana.edu)

In speech perception and spoken word recognition, a lexical "neighbor" of a target word is a phonetically similar word that competes with the target during lexical access. A commonly-used definition is that neighbors differ from one another by a single segment. While this basic definition accounts for many results of lexical decision and word recognition tasks among monosyllabic words, it also suggests that approximately half of the words in the English lexicon are hermits (words without neighbors). To generate a more inclusive definition of neighbor, we created a representative sample of the lexicon (N=1428) to examine spoken word recognition errors, and we used the incorrectly selected competitors as a means of indexing "neighbor" status. Analysis of over 15,000 errors reveals several patterns: (1) Errors tend to be of higher frequency than the target word; (2) Less than 1/3 of the errors were neighbors according to the traditional definition of neighborhood density; and (3) Errors differed in phonetic similarity more as the S/N ratio decreased. We use these data to test several alternative definitions of "neighbor" and their ability to predict actual word recognition errors. [Work supported by NIH-NIDCD R01 00111 and T32 00012]

2pSCc34. Articulatory constraints on coarticulation in fricatives as a function of tongue involvement.

Carole E. Gelfer (William Paterson University, 300 Pompton Road, Wayne, NJ 07470, USA, gelferc@wpunj.edu), Fredericka Bell-Berti (St. John's University, 8000 Utopia Parkway, Queens, NY 11439, USA, bellf@stjohns.edu)

Numerous studies have shown coarticulatory effects of vowel rounding on the spectra of sibilants. However, the extent to which the tongue contributes to coarticulation is not clear because most studies have used only fricatives whose constriction is formed with the tongue. This study examines the effects of vowel environment on coarticulation in fricatives under two conditions: when the constriction is formed with the tongue (i.e., [s]) and when it is not (i.e., [f]), paired with the English vowels occurring at the extremes of the vowel quadrilateral. Preliminary results for adult native speakers of American English suggest that in the absence of tongue constraint (i.e., [f]), there is a greater effect of vowel context on the fricative spectrum. This result reinforces the importance of considering the freedom of an articulator to anticipate an upcoming gesture when examining the extent of coarticulation.

2pSCc35. Acquisition of rhythm: evidence from spontaneous L2 speech.

Irina Shport (University of Oregon, Department of Linguistics, 1290, Eugene, OR 97403, USA, ishport@uoregon.edu)

This study examined the acquisition of Japanese rhythm by comparing the rhythmic characteristics of elicited spontaneous speech of English learners in the setting of an oral proficiency test. Two metrics for the quantification of speech rhythm that have been shown to reliably discriminate rhythmic classes in both L1 and L2 speech (White and Mattys, 2007) were utilized for the analysis - 1) the proportion of vowel intervals within the sentence (%V), and 2) the standard deviation of vowel intervals divided by the

mean vowel interval duration and multiplied by 100 (VarcoV). A difference in the rhythm scores of novice-mid and intermediate-high learners of Japanese was found. Furthermore, the learners' scores were significantly different from Japanese native speakers' scores, suggesting that speakers start acquiring speech rhythm relatively early, but they are still short of the target at the intermediate level of proficiency. By comparing the obtained scores with scores of stress-timed languages, such as English, the difference in the scores may be attributed to an L1 transfer into the rhythmic organization of the L2 speech.

2pSCc36. Effects of training, style, and rate of speaking on speech perception of young people in reverberation. Nao Hodoshima (Dept. of Electrical and Electronics Engineering, Sophia University, 7-1 Kiyoi-cho, Chiyoda-ku, 102-8554 Tokyo, Japan, n-hodosh@sophia.ac.jp), Takayuki Arai (Dept. of Electrical and Electronics Engineering, Sophia University, 7-1 Kiyoi-cho, Chiyoda-ku, 102-8554 Tokyo, Japan, arai@sophia.ac.jp), Kiyohiro Kurisu (TOA Corporation, 2-1 Takamatsusho, Takarazuka, 665-0043 Hyogo, Japan, kurisu_kiyohiro@toa.co.jp)

Because of the difficulty of listening to speech in reverberation (e.g. train stations or airports), we need to find characteristics of intelligible speech sounds that are appropriate for announcements over loudspeakers in public spaces. This study investigated the effects of training (seven talkers who have received speech training or not), style (conversational/clear) and rate (normal/slow) of speaking on speech perception of young people in simulated reverberant environments. The talkers were instructed to speak nonsense words embedded within a carrier sentence clearly or normally in an anechoic room, and listening tests were carried out with young people in simulated reverberant environments. Results showed that correct rates significantly differed among the talkers, but no difference in correct rates was found between the two speaking rates, and conversational speech had significantly higher correct rates than clear speech. Casual inspections of the stimuli indicate that vowels are enhanced as well as consonants in clear speech so that clear speech had lower correct rates than conversational speech due to increased reverberant masking in clear speech compared to conversational speech. [Work supported by Sophia University Open Research Center from MEXT.]

2pSCc37. Prosody in French theatrical declamation traced backwards in time. George Brock-Nannestad (Patent Tactics, Resedavej 40, DK-2820 Gentofte, Denmark, pattac@image.dk)

Listening to early recordings of a French theatrical personality, such as Sarah Bernhardt, the emphatic emission of words is very noticeable. It is sometimes ascribed to Bernhardt's theatrical mannerisms, but possibly also caused by an erroneous speed of reproduction. But even at correct speed, huge variations in intenseness are very noticeable. Literature does not give a clear indication whether this manner of performance was considered deviating at the time (ca. 1900). New research into the original handwritten patent documents of Édouard-Léon Scott (filed 1857 and 1859) has analyzed his method of recording the performance of the human voice. Scott documented the performance of a complete stanza from a French play, and looking at his graphical output and interpretation it becomes clear that the prosody of Mme. Bernhardt was already in place in 1859. Having identified the French play, even earlier descriptions of the prosodic approach in French declamation were found. The paper also discusses the patents of Édouard-Léon Scott and his desire for the use of his invention, which was different from that of his manufacturer, Rudolph Koenig. Furthermore, the prosodic documentation will be demonstrated, and a historically informed reproduction of a recording by Sarah Bernhardt will be presented.

2pSCc38. Articulatory coordination in German and English onset and coda clusters. Marianne Pouplier (Institute of Phonetics and Speech Processing, University of Munich, 80799 Munich, Germany, pouplier@phonetik.uni-muenchen.de), Stefania Marin (Institute of Phonetics and Speech Processing, University of Munich, 80799 Munich, Germany, stefania.marin@gmail.com)

Articulatorily, onset consonants differ from coda consonants in their spatial as well as their temporal characteristics in that onsets exhibit greater spatial displacement and are generally less variable than codas. As to their

temporal organization, it has been claimed that onsets, but not codas exhibit a so-called c-center organization in which all consonant gestures of a cluster are organized as a single complex structure with respect to the vowel, while consonant clusters in codas have been hypothesized to be organized sequentially [Browman & Goldstein 1988 *Phonetica*]. While some studies have confirmed the c-center hypothesis for onsets but not codas, others found a c-center organization for both onset and coda consonants. The current study uses German and English onset and coda clusters to test systematically the c-center hypothesis on the basis of EMA data. Results for both languages indicate that a c-center organization may be observed in both onset and coda position, yet it does not appear consistently for all clusters. Also cross word-boundary clusters may exhibit c-center organization, possibly indicating resyllabification.

2pSCc39. Acoustic consequences of gestural intrusion errors. Stefania Marin (Institute of Phonetics and Speech Processing, University of Munich, 80799 Munich, Germany, stefania.marin@gmail.com), Marianne Pouplier (Institute of Phonetics and Speech Processing, University of Munich, 80799 Munich, Germany, pouplier@phonetik.uni-muenchen.de), Jonathan Harrington (Institute of Phonetics and Speech Processing, University of Munich, 80799 Munich, Germany, jmh@phonetik.uni-muenchen.de), Susanne Waltd (IPS, LMU, Schellingstr. 3/II, 80799 München, Germany, susanne@phonetik.uni-muenchen.de)

An increasing number of studies have elicited speech errors in articulatory experiments, and thus evaluated errors on the basis of their articulatory properties, yet there has been no study that relates the articulatory variability observed in errors to the acoustic consequences. The present study aims to evaluate the acoustic properties of errorful/error-free [k] and [t] stimuli with known articulatory configurations (and articulatorily categorized as errorful/error-free), and determine the extent to which articulatory errors have acoustically observable consequences. The most frequent error observed in the articulatory data is an intrusion of a second gesture, resulting in a simultaneous presence of both tongue tip and tongue body gestures during production of [k] or [t]. Spectral analysis of these stimuli's bursts shows that [k] and [t] are differently affected by intrusion gestures: intrusion errors during intended [k] tokens are more robustly grouped with error-free [k] across speakers, while intrusion errors during intended [t] tokens are more variable, some being acoustically like [t], some like [k], and some like neither category, matching previous perceptual results (Pouplier & Goldstein 2004, *Journal of Phonetics*). These results will be discussed in the context of different timing patterns between the two co-occurring gestures during intrusion errors.

2pSCc40. Embodiment, not imitation, leads to the replication of timing phenomena. Piers Messum (112 Warner Road, SE5 9HQ London, UK, p.messum@gmail.com)

In many languages, a spoken vowel is shorter before phonologically voiceless consonants than before voiced ones. In West Germanic languages, including English, tense and lax vowels have different lengths and there is a characteristic 'stress-timed' rhythm. For these phenomena and some others, it is generally assumed that time is the controlled variable for production, and that children replicate these speech behaviors through the reproduction of timing patterns abstracted from the adult input (i.e. by imitation). The developmental data does not support these assumptions, and an imitative mechanism would present a young speaker with a highly complex challenge. Instead, these behaviors are more plausibly the result of the embodiment of speech. Embodiment goes beyond laryngeal and articulatory function. A child's speech breathing is not a scaled-down version of the adult model but a distinctly different skill, and one that must be learnt during speech production. Similarly, the aerodynamic setting of child speech differs significantly from that of adults. The constraints that these factors impose become manifest in speech as changes in timing, but these changes are epiphenomenal, not modelled directly. Phenomena particular to West Germanic languages reflect the style of speech breathing these languages require of a child.

2p TUE. PM