

Session 4pSW

Speech Workshop: Cross Language Speech Perception and Linguistic Experience: Poster Session A

Linda Polka, Chair

*School of Communication Sciences and Disorders, McGill Univ., Montreal, QC H3G 1A8, Canada***Contributed Papers**

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

4pSW1. Acoustic correlates of Tone 3 and Tone 4 in Mandarin. Chia-Hsin Yeh (1627 Spartan Village Apt. L, East Lansing, MI 48823)

Perceptual similarity between Tone 2 and Tone 3 in Mandarin was widely discussed in previous studies [Moore and Jongman (1997), Huang (2004), Bent (2005)]. Other tonal contrasts are hardly addressed. However, recent findings of Mandarin tones show that Tone 3 and Tone 4 are confusing in terms of descent slope. A big difference between previous studies and current ones is kinds of Tone 3 stimuli: previous studies used isolated Tone 3, namely, 214, and current studies used derived Tone 3, namely, 21, which is coarticulated with other tones. Spectrographic analysis was conducted. Descent slope and height of starting pitch were found critical between Tone 4 and derived Tone 3. Then three continua with ten instances for each were synthesized in respect of three different factors, descent slope, height of starting pitch, and height of ending pitch, which makes 30 stimuli total in three experiments, an AXB discrimination test, an identification test, and a lexical decision test. Results show that starting pitch height is a salient acoustic factor for 30 Mandarin natives, although descent slope was predicted more important. In addition, lexicon has an influence on ambiguous instances between endpoints of Tone 3 and Tone 4. Current findings may make contributions to understanding of perception, articulation, and lexicon during speech processing.

4pSW2. Intelligibility of foreign-accented speech in noise for younger and older adults. Sarah Hargus Ferguson (Dept. of Speech-Lang.-Hearing: Sci. and Disord., Univ. of Kansas, Dole Ctr., 1000 Sunnyside Ave., Rm. 3001, Lawrence, KS 66045, safergus@ku.edu), Kyung Ae Keum, Allard Jongman, and Joan A. Sereno (Univ. of Kansas, Lawrence, KS 66045)

In a previous experiment [Ferguson *et al.*, *J. Acoust. Soc. Am.* **118**, 1932 (2005)], young listeners with normal hearing, older adults with essentially normal hearing, and older adults with hearing loss were found to be similarly affected by the presence of a foreign accent on a word identification task in various listening conditions. This result stands in sharp contrast with an extensive literature demonstrating that the negative effect of speech distortions such as noise, reverberation, and time compression is disproportionately great for older versus younger adults. The conclusions of the previous study, however, were tempered by the presence of floor effects in the data for the older adults with hearing loss identifying accented speech in noise (12-talker babble at a signal-to-babble ratio of +3 dB). The present experiment extended the earlier one by presenting the same monosyllabic word materials to new groups of young normal, elderly normal, and elderly hearing-impaired listeners in three new signal-to-babble ratios: +6 dB, +9 dB, and +12 dB. The results will be discussed in terms of their implications for older adults interacting with foreign-accented speakers in healthcare settings. [Project supported by University of Kansas General Research Fund.]

4pSW3. The acquisition of sonorant voicing. Jagoda Sieczkowska (Inst. of Nat. Lang. Processing, Univ. of Stuttgart, Azenbergstr. 12 D-70174 Stuttgart, Germany, jagoda.sieczkowska@ims.uni-stuttgart.de)

In the view of Exemplar Theory phonetic variants of words are stored by speakers in their memory as a set of exemplars organized into clusters and

grow stronger or decay over time depending on language experience (Bybee, 2002). In Polish and French sonorant consonants in clusters are devoiced word finally, which does not occur in German and English because of phonotactical constraints. A study based on automatic speech alignment and voicing profile extraction from Polish, French, German, and English speech corpora aims at defining voicing sonorant temporal structure and voicing profile acquisition. The hypothesis is that Polish and French native speakers transfer their L1 exemplars during L2 acquisition of English and German and thus devoice sonorant. After extracting voicing profiles from recordings of native Polish and French speakers speaking English and German as an L2/L3, comparison of voicing profiles of English and German native speakers will be made in order to define the exemplar transfer.

4pSW4. Perceptual adaptation to sinewave-vocoded speech across languages. Tessa Bent (Dept. of Speech and Hearing Sci., Indiana Univ., 200 S. Jordan Ave., Bloomington, IN 47405, tbent@indiana.edu), Jeremy L. Loebach (Macalester College, Saint Paul, MN 55105), Lawrence Phillips, and David B. Pisoni (Indiana Univ., Bloomington, IN 47405)

Listeners rapidly adapt to many forms of degraded speech. What level of information drives this adaptation (e.g., acoustic, phonetic, lexical, syntactic), however, remains an open question. In the present study, three groups of listeners were passively exposed to sinewave-vocoded speech in one of three languages (English, German, or Mandarin) to manipulate the level(s) of information shared between the training languages and testing language (English). Two additional groups were also included to control for procedural learning effects. One control group received no training, while the other control group was trained with spectrally rotated English materials. After training, all listeners transcribed eight-channel sinewave-vocoded English sentences. The results demonstrated that listeners exposed to German materials performed equivalently to listeners exposed to English materials. However, listeners exposed to Mandarin materials showed an intermediate level of performance; their scores were not significantly different from the German or English groups but were also not significantly different from the two control groups. These results suggest that lexical access is not necessary for perceptual adaptation to degraded speech, but rather similar phonetic structure between the training and testing languages underlies adaptation to degraded speech. [Work supported by NIH-NIDCD T32 DC-00012 and R01 DC-000111.]

4pSW5. Assessment of perceptual assimilation before and after training. James D. Harnsberger (Dept. of Commun. Sci. and Disord., Univ. of Florida, Gainesville, FL 32611, jharns@ufl.edu)

Many non-native speech sounds are challenging to perceive and, ultimately, to acquire. The development of a model of cross-language speech perception and learning has been hampered by, among other issues, the high variability commonly observed in the perceptual assimilation of non-native contrasts due to phonetic context, talker, and a variety of indexical properties of speech. This variability may reflect persistent patterns in perception that influence learning, or it may represent only an early sensitivity to low-level phonetic detail that is attenuated as experience increases. To examine this issue, the perceptual assimilation by native speakers of American English of a variety of voicing and place contrasts from Hindi was assessed be-

fore and after training in a paired-associate word learning task. Training utilized tokens from four of the six talkers used in perceptual assimilation tasks. The purpose of the study was to determine whether or not assimilation patterns could be greatly modified and/or simplified over the course of word learning in the laboratory. If limited training resulted in large changes in perceptual assimilations, then models of cross-language speech perception and learning will need to rely on assimilation data elicited via procedures that result in stable and repeatable patterns.

4pSW6. Production of disyllabic Mandarin tones by children. Pui-san Wong (Dept. of Speech Commun. Arts and Sci., Brooklyn College, CUNY, 2900 Bedford Ave., Brooklyn, NY 11210, pswRes.@gmail.com) and Wini-ferd Strange (CUNY Grad Ctr., New York, NY 10016)

Two- to six-year-old children's productions of Mandarin lexical tones in familiar disyllabic words were examined to determine the time course of tone development and the effect of complexity of fundamental frequency contours (f₀) on rate of acquisition. Disyllabic tone (DT) productions of 44 children and 12 of their mothers were elicited in a picture naming task and were low-pass filtered to eliminate lexical information. Judges categorized the DTs based on the filtered stimuli without lexical support. Adults' and children's DTs were categorized with 96% and 65% accuracy, respectively. Tone accuracy increased while intersubject variability decreased with age. Children as old as 6 years did not attain adultlike accuracy. Different DT combinations developed at different rates. DTs with more complex f₀ contours were more difficult for children. Substitution patterns and acoustic analysis revealed that when producing tones with large transitions at the syllable boundary, children tended to modify the f₀ in the first syllable to reduce the f₀ shift at the boundary of the first and second syllable, resulting in more errors in the tones in the first syllable than in the second syllable. The results suggest physiological constraints on tone development. [Work supported by NIH 5F31DC8470 and NSF.]

4pSW7. Effectiveness of a perceptually weighted measure on automatic evaluation of a foreign accent in a second language. Hiroaki Kato (NICT/ATR, 2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan, kato@atr.jp), Shizuka Nakamura (GITI, Waseda Univ., Shinjuku-ku, Tokyo 169-0051, Japan), Shigeki Matsuda (NICT/ATR), Minoru Tsuzaki (Kyoto City Univ. Arts, Nishikyo-ku, Kyoto 610-1197, Japan), and Yoshinori Sagisaka (GITI, Waseda Univ./NICT, Tokyo, Japan)

An empirical study is carried out to achieve a computer-based methodology for evaluating a speaker's accent in a second language as an alternative to a native-speaker tutor. Its primary target is the disfluency in the temporal aspects of an English learner's speech. Conventional approaches commonly use measures based solely on the acoustic features of given speech, such as segmental duration differences between learners and reference native speakers. However, our auditory system, unlike a microphone, is not transparent: it does not send incoming acoustic signals into the brain without any treatment. Therefore, this study uses auditory perceptual characteristics as weighting factors on the conventional measure. These are the loudness of the corresponding speech segment and the magnitude of the jump in loudness between this target segment and each of the two adjacent speech segments. These factors were originally found through general psychoacoustical procedures [H. Kato et al., *JASA*, **101**, 2311–2322 (1997); **104**, 540–549 (1998); **111**, 387–400 (2002)], so they are applicable to any speech signal despite the difference in language. Experiments show that these weightings dramatically improve evaluation performance. The contribution of psychoacoustics to evaluation methodology of second language speech is also discussed. [Work supported by RISE project, Waseda Univ. and KAKENHI 20300069, JSPS.]

4pSW8. Effects of “catch” trials on the discrimination of American English vowels by native Japanese speakers. Takeshi Nozawa (Lang. Education Cntr., Faculty of Economics, Ritsumeikan Univ. 1-1-1 Nojihigashi, Kusatsu, Shiga 525-8577, Japan)

Native Japanese speakers' ability to discriminate American English vowels was tested in two different formats, in both of which they heard three stimuli per trial. The three stimuli in each trial were produced by three different talkers. In AXB format, subjects were to decide if the second stimulus included categorically the same vowel as the first or the third stimulus. In other words, the subjects knew that one of the three stimuli was categori-

cally different from the other two. In the other format, the subjects heard three stimuli in each trial, and chose one stimulus that includes a categorically different vowel from the other two. Each vowel contrast was tested by different trials which contained an odd item out, and catch trials which contained three tokens each from each of the vowel contrast pairs. Subjects were visually presented with four boxes on a computer screen with the numbers 1, 2, and 3, and “No.” The “No” button was provided when no differences were detected. The results reveal that with catch trials included, discriminating American English vowels is a lot more difficult. The results imply the Japanese subjects' susceptibility to talker variance and inability to discriminate American English vowels categorically.

4pSW9. A twin study of speech perception and production in first and second language among Chinese children learning English as a second language. Simpson Wai-Lap Wong, Dorothy Bishop (Dept. of Experimental Psych., Univ. of Oxford, S. Parks Rd., Oxford, OX1 3UD, UK, wls Wong@gmail.com), and Connie Suk-Han Ho (Univ. of Hong Kong, Hong Kong)

The faculty of language processing in our brain develops with the constraint of our genetic dispositions and also our experiences. Do the same genetic influences affect learning of phonology in L1 and L2? This study examines the genetic and environmental effects on the individual differences on L1 and L2 speech perception and production among Chinese children learning English as a second language. By employing a twin-study design, 150 pairs of monozygotic and 150 pairs of same-sex dizygotic twins aged from 4 to 11 were tested. Children's speech perception and production in both languages were assessed with an AXB speech perception task of minimal word pairs and two picture naming tasks, respectively. Children's non-verbal IQ was also measured. Scores of accuracy will be computed for each task. Analyses will be conducted to consider the relationship between accuracy in perception and production in L1 and L2, to estimate the extent of genetic contribution to speech perception, and to determine whether heritable individual differences are accounted for by a common factor, or whether different factors influence proficiency in L1 and L2.

4pSW10. Lexical frequency, orthographic information, and first-language effects on second-language pronunciation. Ron I. Thomson (Dept. of Appl. Linguist., Brock Univ., 500 Glenridge Ave., St. Catharines, ON L2S 3A1, Canada, rthomson@brocku.ca) and Talia Isaacs (McGill Univ., Montreal, QC H3A 1Y2, Canada)

In L2 speech learning, lexical frequency may play a facilitative role, whereby perception and production of sounds found in high-frequency lexical items will develop before the perception and production of the same categories found in low-frequency lexical items (see Munro and Derwing, 2008). Orthographic information may also facilitate learning by disambiguating L2 sounds (see Erdener and Burnham, 2005), particularly in known words. This study examines the role of lexical frequency, orthographic information, and a learner's L1 in the development of L2 speech perception and production. Thirty-eight Mandarin and Slavic participants were asked to repeat a word list comprising ten target English vowels, each embedded in three separate monosyllabic verbs and varying in lexical frequency. Recordings of the L2 productions were obtained under three counter-balanced conditions: (1) after hearing an auditory prompt accompanied by the written form of the word; (2) after hearing an auditory prompt with no written form provided; and (3) with no auditory prompt but the written form provided. To measure L2 performance, L1 English listeners were asked to identify the vowel they perceived in each recorded item. Results were examined to determine what influence lexical frequency and orthographic information might have had on L2 performance.

4pSW11. A closer look at perceptual epenthesis in cross-language perception. Lisa Davidson and Jason Shaw (Dept. of Linguist., New York Univ., 726 Broadway, New York, NY 10003, lisa.davidson@nyu.edu)

Previous studies have shown that listeners have difficulty distinguishing non-native CC sequences from their CVC counterparts (e.g. Japanese [ebzo]-[ebuzo], English [zgam0]-[zəgam0]). Some have argued that the phonology “repairs” non-native consonant clusters with vowel insertion (“perceptual epenthesis”), causing listeners to respond that CC and CVC sequences are the same. Production studies, however, have shown that speak-

ers produce non-native CC sequences many ways, including epenthesis, changing C1, deleting C1, and prothesis. To test whether these other repairs of non-native clusters are also difficult to distinguish, English listeners ($n=37$) participated in an AX discrimination task that paired C1CVCV stimuli with CəCVCV (epenthesis), əCVCVCV (prothesis), CVCV (deletion), or C2CVCV (C1 change). Each of these repairs caused some difficulty and there was an interaction between repair type and manner combination (stop-stop, stop-nasal, fricative-stop, fricative-nasal). Listeners were more accurate when C1 was deleted (81% accurate), followed by epenthetic and C1 changes (both 77%), then by prothesis (59%). Furthermore, compared to previous studies testing only the epenthetic repair, presenting listeners with various repair types dramatically improves discrimination in the epenthesis condition. These results indicate that perceptual epenthesis may be a task effect and that top-down phonological influences are more complex than previously assumed. [Research supported by NSF.]

4pSW12. Native and non-native clear speech production. Rajka Smiljanic (Linguist., Univ. of Texas at Austin, 1 University Station B5100, Austin, TX 78712-0198) and Ann Bradlow (Northwestern Univ., Evanston, IL 60208)

Previous work established that both native and non-native clear speech increase intelligibility for native and proficient non-native listeners. However, non-native clear speech enhanced intelligibility less than native clear speech. In this study, we examine native and non-native conversational and clear speech productions with an eye on the differences in hyperarticulation strategies that may underlie a smaller clear speech intelligibility benefit. Results revealed a slower conversational speaking rate for non-native talkers compared with native talkers. Clear speaking rate was similar for both groups suggesting a limit in the speaking rate decrease in intelligibility enhancing clear speech. The durational distinction between tense and lax vowels was increased less in non-native speech due to the smaller speaking rate decrease in clear speech. Analyses of the stop voice onset time (VOT) showed that non-native talkers lengthened the voiced category in the negative VOT direction, while the voiceless category was lengthened less in clear speech. Finally, examination of vowel space expansion in clear speech revealed differences between the two talker groups reflecting non-native talkers' inexperience in producing some of the vowel targets. Combined, these results suggest that a complex interaction of articulatory difficulty, proficiency, and native language background determines levels of non-native clear speech intelligibility.

4pSW13. Measuring sensitivity to short-term deviations in acoustic cue correlation. Kaori Idemaru (Dept. of East Asian Lang. and Lit.s, Univ. of Oregon, Eugene, OR 97403) and Lori L. Holt (Carnegie Mellon Univ., Pittsburgh, PA 15213)

Listeners are sensitive to correlations among the multiple probabilistic acoustic cues that define speech categories. In English stop consonant productions, for example, fundamental frequency (f_0) is strongly correlated with voicing. Reflecting this regularity, perception of syllables varying in voice onset time is shifted with changes in f_0 . Such sensitivity to the long-term regularities of a language must be balanced with enough flexibility that speech perception is able to accommodate deviations from these regularities, such as those that may arise variations from idiolect, dialect, or accented speech. The present experiments manipulate short-term acoustic cue correlations experienced in online perception to investigate the interplay between sensitivity to long-term and short-term acoustic cue correlation. Using overt categorization and eyetracking, the present results indicate that speech categorization is influenced by local shifts in acoustic cue correlations that deviate from long-term regularities of English. These experiments further examine the time course of this short-term learning and the degree to which it generalizes. The data suggest that listeners continually monitor speech input for regularity and tune online speech perception in relation to these regularities.

4pSW14. Articulatory analysis of foreign-accented speech using real-time MRI. Joseph Tepperman (Viterbi School of Eng., Univ. of Southern California, 3740 McClintock Ave., EEB 400, Los Angeles, CA 90089, tepperma@usc.edu), Erik Bresch, Yoon-Chul Kim, Louis Goldstein, Dani Byrd, Krishna Nayak, and Shrikanth Narayanan (Univ. of Southern California, Los Angeles, CA 90089)

We present the first study of nonnative English speech using real-time MRI analysis. The purpose of this study is to investigate the articulatory nature of “phonological transfer”—a speaker’s systematic use of sounds from their native language (L1) when they are speaking a foreign language (L2). When a non-native speaker is prompted to produce a phoneme that does not exist in their L1, we hypothesize that their articulation of that phoneme will be colored by that of the “closest” phoneme in their L1’s set, possibly to the point of substitution. With data from three native German speakers and three reference native English speakers, we compare articulation of read phoneme targets well documented as “difficult” for German speakers of English (/w/ and /dh/) with their most common substitutions (/v/ and /d/, respectively). Tracking of vocal tract organs in the MRI images reveals that the acoustic variability in a foreign accent can indeed be ascribed to the subtle articulatory influence of these close substitutions. This suggests that studies in automatic pronunciation evaluation can benefit from the use of articulatory rather than phoneme-level acoustic models. [Work supported by NIH.]

4pSW15. The perception of speech and nonspeech analogs by first- and second-language speakers. Paul Iverson, Anita Wagner, and Stuart Rosen (Res. Dept. Speech Hearing and Phonetic Sci., University College London, 2 Wakefield St., London WC1N 1PF, UK)

There are marked differences in how native speakers of Sinhala and English perceive the English /w/-/v/ distinction; Sinhala speakers who have learned English as a second language are typically near chance at identification and have less than half the discrimination sensitivity of native English speakers. This poor discrimination ability is remarkable because the acoustic cues for /w/ and /v/ are varied, being contrastive in formant frequencies, amount of friction, and amplitude envelopes. The present project explored these differences in auditory sensitivity by manipulating speech and nonspeech stimuli, with the aim of investigating how close an auditory stimulus needs to be to natural speech in order for cross-language perceptual differences to emerge. Synthesized VCV speech stimuli were created to model natural recordings. Nonspeech stimuli were created by removing dynamics from the stimuli (e.g., flat pitch and amplitude, no formant movement), producing a “buzz” that was acoustically similar to the speech stimuli. Discrimination results for these stimuli will be reported for native speakers of Sinhala and English, in order to evaluate whether cross-language specialization for speech may occur at a precategorical auditory-phonetic level of processing, or whether specialization is contingent on the stimuli being perceived as intelligible speech.

4pSW16. Proficiency instrumentation for cross language perception studies. Jared Bernstein (Pearson, 299 S. California Ave., Ste. 300, Palo Alto, CA 94306, Jared.Bernstein@Pearson.com)

Studies in cross language speech perception sometimes use self-reported values of language proficiency as an independent variable. In relation to other measurement methods, proficiency self-reports have been found susceptible to several sources of bias including demographic variables and proficiency level itself [S. Yamashita, *Six Measures of JSL Pragmatics*, U. Hawaii Press (1996)]. An inaccurate scaling of an independent variable can limit the precision of performance models that are based on that variable’s relation to the dependent variable. This paper presents a convenient method (Versant tests) for measuring spoken language proficiency that locates subjects on a 60 point scale with a standard error measurement of less than 3 points, thus reliably locating subjects into 20 levels along a linear continuum of proficiency. The testing procedure takes about 15 minutes to complete. The test-retest reliability of Versant scores for spoken English and Spanish are 0.97 and 0.96, respectively, both with correlations greater than 0.85 with independent human rated tests. The development, operation, and psychometric properties of the Versant tests are presented briefly, along with examples of their use in evaluating language instruction programs and standard listening comprehension tests.

4pSW17. Why do adults vary in how categorically they rate the accuracy of children’s speech? Eden Kaiser (Prog. in Linguistic, Univ. of Minnesota, 214 Nolte Ctr., 315 Pillsbury Dr. SE, Minneapolis, MN 55455, kaise113@umn.edu), Benjamin Munson (Univ. of Minnesota, Minneapolis, MN 55455), Fangfang Li (Univ. of Lethbridge, Lethbridge, AB T1K 3M4, Canada), Jeff Holliday (Ohio State Univ., Columbus, OH 43210), Mary Beckman (Ohio State Univ., Columbus, OH 43210), Jan Edwards, and Sarah Schellinger (Univ. of Wisconsin-Madison, Madison, WI 53706)

In a recent experiment using continuous visual analog scales (VAS) to examine adults' perception of children's speech accuracy, listeners varied in the extent to which they categorically perceived children's English and non-English productions [Munson *et al.*, American Speech-Language-Hearing Association (2008)]. Some listeners utilized all points on the VAS line for their responses, while others grouped responses around discrete locations on the line. It is hypothesized that differences in categoricity of responses across listeners might relate to listeners attending to either categorical linguistic information (i.e., identifying phonemes in a word, which would promote more categorical labeling) or gradient indexical information (i.e., identifying the child's sex or age, which would promote more continuous labeling). If this is true, it should be possible to elicit differences in categoricity of fricative "goodness" judgments in individual listeners by priming them to listen to linguistic variables (by interleaving fricative judgment trials with trials in which they categorize the vowel spoken by the child) or indexical variables (by interleaving fricative judgment trials with trials in which they identify the child's sex). This paper reports on an experiment designed to test this. Results will help us better understand individual response patterns in cross-language speech perception experiments.

4pSW18. Perceptual studies of two phonological voicing patterns. Scott Myers (Dept. of Linguist., 1 Univ. Sta. B5100, Univ. of Texas, Austin, TX 78712, s.myers@mail.utexas.edu)

Two of the most common phonological patterns in the world's languages are final devoicing and regressive voicing assimilation. In final devoicing, there is a contrast in obstruent voicing generally, but in word-, phrase-, or syllable-final position, only voiceless obstruents are allowed (e.g., Dutch, Russian, Wallon). In regressive voicing assimilation, the first of a sequence of two obstruents is required to agree with the second in voicing (e.g., Russian, Catalan, Hebrew). These phonological patterns can be related to patterns of speech production: anticipation of the spread glottis posture in pause and laryngeal coarticulation within consonant clusters. The studies reported here investigate the perceptual consequences of these coarticulatory effects. Native speakers of English listened to words contrasting in the voicing of the final consonant (cease/seize) and identified which word they heard. The words were excised from utterance-final position or from phrase-medial position preceding a vowel, nasal, a voiced obstruent, or a voiceless obstruent. The subjects averaged about 90 percent correct identification. But they still displayed a significant bias toward identifying as voiceless the final consonant in utterance-final words as well as in words preceding a voiceless obstruent. This pattern of identification errors could be the diachronic basis of the phonological patterns.

4pSW19. The relationship between speech perception and production in second language learners. Bing Cheng (English Dept., Inst. of Intl. Studies, Xi'an Jiaotong Univ., Xi'an, China, 710049) and Yang Zhang (Univ of Minnesota, Minneapolis, MN 55455)

This study aims to examine the relationship between perception and production skills in second language learners. Thirty-nine Chinese college students who have received at least eight years of English-as-a-second-language (ESL) education in school participated in the study. The data were collected using two programs: HearSay from Communication Disorders Technology Inc., Indiana, and a perceptual test program developed by Yang Zhang at the University of Minnesota. Digital recordings of each participant's production of the individual words used in the perceptual test program were rated in the scale of 1 to 5 by two native speakers of English. Results demonstrated the existence of complex relationships between perception and production in ESL learners. Overall, the consonants demonstrated a significant positive correlation between perception and production ($p < 0.01$) with voicing showing strongest effect. Vowels, on the other hand, did not show significant correlations in any of the five categories examined: tense/lax, high/low, front/back, diphthong, and vowel insertions. Four patterns were identified for the individual phonemes: (a) good perception and good production, (b) poor perception and poor production, (c) good percep-

tion and poor production, and (d) poor perception and good production. The results are discussed with respect to speech learning models and practical instructional approaches.

4pSW20. Indexical perception in Chinese: The influence of first-language (L1) Taiwanese on second-language (L2) Mandarin. Yufen Chang (2451 E. 10th St., Apt. 909, Bloomington, IN 47408, yufchang@indiana.edu)

A pilot study on indexical perception in Chinese hypothesized that L1 Taiwanese speakers' L2 Mandarin production could be detected as having a nonnative accent. If the nonnative accent was detectable, what phonetic properties (the phonetics of consonants or vowels, or the prosody) did L1 Taiwanese speakers' Mandarin exhibit such that these attributes entailed indexical information? Additionally, the study argued that L1 Mandarin speakers would show higher sensitivity to the nonnative accent than L1 Taiwanese speakers. The results have shown that only in sentence production, the sentences being either low-pass filtered or unfiltered, could the L1 Taiwanese talkers be possibly indexed correctly. It was also found that the listeners perceived more accurately when they encountered retroflex stimuli in monosyllabic word condition. Based on these findings, it was concluded that the prosody and the retroflex consonants seemed to reveal much of the talker's L1 information. However, concerning the hypothesis that L1 Mandarin speakers were more sensitive to the nonnative accent, the results did not support this. The findings in this pilot study are preliminary, given the fact that the corpus was small. A bigger-scale study is being carried out to gain more insights into indexical perception in Chinese.

4pSW21. Automatic auditory discrimination of vowels in simultaneous bilingual and monolingual speakers as measured by the mismatch negativity (MMN). Monika Molnar, Shari Baum, Linda Polka (McGill Univ., 1266 Pine Ave. W, Montreal, H3G 1A8 Canada, monika.molnar@mcgill.ca), Lucie Menard (Universite du PQ à Montréal, Montreal, Canada), and Karsten Steinhauer (McGill Univ., Montreal, H3G 1A8 Canada)

MMN responses reflect whether language users have developed long-term memory traces in response to phonemes and whether they are able to perceive small acoustic changes within speech sound categories. Subtle acoustic changes within phonemes are often irrelevant to monolingual perceivers, but can be crucial for bilingual perceivers if the acoustic change differentiates the phonemes of their two languages. In the present study, we investigated whether bilinguals are sensitive to such acoustic changes. We recorded MMN responses from monolingual (English, French) and simultaneous bilingual (English/French) adults using an auditory oddball paradigm in response to four vowels: English [u], French [u], French [y], and an acoustically-distinct (control) [y]. In line with previous findings, monolinguals were more sensitive to the phonemic status of the vowels than to the acoustic properties differentiating the sounds. Bilingual speakers revealed a different pattern: they demonstrated overall slower discrimination responses to all sounds, but showed almost equal sensitivity to phonemic and phonetic/acoustic differences. The results suggest that bilingual speakers exhibit a more flexible but less uniquely-specified perceptual pattern compared to monolingual speakers.

4pSW22. Comparison between Japanese children and adult's perception of prosodic politeness expressions. Takaaki Shochi (Div. of Cognit. Psych., Kumamoto Univ., Kurokami 2-40-1, Kumamoto City, 860-8555, Japan, shochi38@gmail.com), Donna Erickson (Showa Univ. of Music, Kawasaki-city, Kanagawa pref., 215-0021, Japan), Albert Riiliard (LIMSI-CNRS, BP 133 F-91403 Orsay cedex, France), Kaoru Sekiyama (Kumamoto Univ., Kurokami 2-40-1, Kumamoto city, 860-8555, Japan), and Veronique Auberge (Domaine Universitaire, BP46 F-38402 Saint Martin d'Heres cedex, France)

Previous work examined the contribution of audio and visual modalities for perception of Japanese social affects by adults. The results showed audio and visual information contribute to the perception of culturally encoded expressions, and show an important synergy when presented together. Multimodal presentation allows foreign adult listeners to recognize culturally encoded expressions of Japanese politeness which they cannot recognize with an audio-only stimuli. This current work analyzes the recognition performance of politeness expressions by Japanese children 13 to 14 years old.

Stimuli, based on one sentence with an affectively neutral meaning, are performed with five different expressions of politeness. Subjects listen three times to each stimulus and judge the intended message of the speaker. The stimuli are presented as audio-only, visual-only, audio-visual. Listeners rate the social status of the hearer and the degree of politeness on a nine-point scale ranging from polite to impolite. The results are analyzed to capture the relative ability of adults and children to use both modalities to recognize social affects. [This work was supported in part by Japanese Ministry of Education, Science, Sport, and Culture, Grant-in-Aid for Scientific Research (C) (2007–2010): 19520371 and SCOPE (071705001) of Ministry of Internal Affairs and Communications (MIC), Japan.]

4pSW23. Do you want to pronounce correctly in a foreign language? Start music lessons! Barbara E. Pastuszek-Lipinska (School of English, Adam Mickiewicz Univ., al. Niepodleglosci 4, 61-874 Poznan, Poland)

Music education is one of the human activities that requires the integration of all human senses and the involvement of all cognitive processes: sensory, perceptual, and cognitive learning, memory, emotion, and auditory and motor processes, music has tended to figure only marginally in an approach to second-language acquisition. In the presentation will be reviewed recent research results on the influence of music lessons on second language speech processing. The presentation will draw a comparison between second language acquisition processes and first language acquisition processes. It will be also provided concise review of the possible explanations of the influence of music on second language speech processing. The review will include acoustic, phonological, neuroscientific, and psycholinguistic approaches.

4pSW24. Mandarin-English bilinguals' accented first-language (L1) vowel production. Haisheng Jiang (School of Lang. and Lit., Univ. of Aberdeen, Aberdeen, AB24 3FX, UK, h.jiang@abdn.ac.uk)

The L1 phonetic categories established in childhood may undergo modification when similar L1 and L2 sounds interact in the process of L2 learning [J. Flege, *Speech perception and linguistic experience: Issues in cross-language research* (1995)]. In this study, Mandarin vowel production by the Mandarin-English bilinguals was presented to Mandarin as well as English listeners for goodness rating. The results showed that both Mandarin-English bilinguals of high L1 use and those of low L1 use differed significantly from Mandarin monolinguals in the production of /y/, a vowel with no counterpart in English. An analysis of inter-speaker variability indicated that some individual Mandarin-English bilinguals, including both speakers of high L1 and low L1 use, were accented in the production of /y/, and /au/. Possible acoustic properties contributing to their accentedness were identified. L2 English learning led to some Mandarin-English bilinguals carrying some English characteristics in their L1 Mandarin vowel production. This study provides further evidence for the claim that the L1 phonetic system established in childhood is susceptible to change. It contributes to the less well-studied field of L2 influence on L1.

4pSW25. Assessing phonetic talent in second language performance: The relationship between perception abilities in the native and the second language and production. Matthias Jilka (Institut of English Linguist., Univ. of Stuttgart, Heilbronner Str. 7, 70174 Stuttgart, Germany, jilka@ifla.uni-stuttgart.de), Giusy Rota (Univ. of Stuttgart, 70174 Stuttgart, Germany), and Susanne Reiterer (Univ. of Tuebingen, 72076 Tuebingen, Germany)

An extensive project that aims to assess innate phonetic talent is introduced. It investigates language performance of 103 native speakers of German in their native language, in English and (to a limited degree) in Hindi. The project's original priority is to establish a talent score that serves as the basis for a neurolinguistic study which attempts to find correlates between phonetic talent and brain anatomy/function. For this purpose tasks investigating neurobiological, psychological, and other factors influencing performance are also conducted. Innate talent is to be distinguished from superficial proficiency by limiting or eliminating the influence of interfering factors such as language experience. The test subjects are investigated with respect to different manifestations of their production (e.g., quasi-spontaneous speech, reading), perception (e.g., discrimination, identification), and imitation abilities in both German and English. A first analysis of the interactions between all perception and production scores

shows a significant correlation at the level of $p < 0.05$. The correlation is especially high ($p < 0.01$) between perception tasks involving the interpretation of intonational features of both English and German and production abilities (again both in German and English), suggesting a certain cross-language predictive power.

4pSW26. Speech perception assessment and training system (SPATS-ESL) for speakers of other languages learning English. James D. Miller, Roy Sillings, Charles S. Watson, and Diane Kewley-Port (Commun. Disord. Technol., Inc., 501 N. Morton St., Sta. 215, Bloomington, IN 47404)

The SPATS software system, originally developed for the hearing-impaired, has been modified for use with ESL learners with TOFEL (pbt) scores near and well above 500. SPATS-ESL includes the identification of syllable constituents: onsets, nuclei, and codas, as well as sentence recognition. The syllable constituent tasks include the progressive introduction of increasing numbers of constituents until the learner becomes adept at the identification of 45 onsets, 28 nuclei, and 36 codas presented by eight talkers in a variety of phonetic contexts. The sentence task emphasizes increasing speed and decreasing errors in the recognition of short, meaningful sentences spoken by a variety of talkers. The sentences are presented in a background of multi-talker babble at five signal-to-noise ratios: +10, +5, 0, -5, and -10 dB. The syllable constituent and sentence tasks are interleaved throughout training. In constituent training, SPATS uses a proprietary training algorithm, Adaptive Item Selection (AIS), which automatically focuses training on individuals' items of intermediate difficulty and is independent of their language history. Proctored tests allow certification of a learner's English speech perception in relation to native-speaker performance.

4pSW27. On-line perception of lexical stress in English and Spanish: The effects of sentence intonation and vowel reduction. Marta Ortega-Llebaria (Dept. of Spanish and Portuguese, Univ. of Texas at Austin, Austin, TX 78712, ortegallebaria@att.net)

Previous research shows that speakers of stress-accent languages rely on pitch-accents to perceive word stress in sentences spoken with declarative intonation, while in unaccented sentences, like post-focal contexts, they rely on other cues, i.e., duration in Spanish or vowel reduction in English. However, there is no experimental evidence on the effect that sentence intonation has in the "on-line" perception of word stress across languages. This experiment examines whether listeners detect word stress faster in unaccented stretches of speech that are preceded by a landmark in the sentence melody, i.e., the focal pitch-accent that always precedes a post-focal sentence, than in unaccented contexts not preceded by a landmark, i.e., reporting clauses. Results show that Spanish listeners identified target words at the beginning and end of post-focal contexts with similar reaction times, while in reporting clauses, target words placed at the end were identified faster than those placed at the beginning. Preliminary results for English show no significant differences. Thus, Spanish listeners re-weight cues to word stress on-line taking into account the patterns of sentence intonation. Sentence melody does not have such a strong effect in English, possibly because vowel reduction is a sufficient cue to an effective perception of word stress.

4pSW28. A cross-language study of compensatory response to formant-shifted feedback. Takashi Mitsuya, Ewen N. MacDonald (Dept. of Psychol., Queen's Univ., Kingston, Canada), David W. Purcell (Univ. Western Ontario, London, Canada), and Kevin G. Munhall (Queen's Univ., Kingston, Canada)

Previous experiments in speech motor learning have demonstrated that the perception of our own voice while we speak plays a role in the control of fundamental and formant frequencies and vocal amplitude. When feedback is changed in real time, subjects alter speech production in attempt to compensate for the perturbation. By testing Japanese talkers in their native and a less familiar language (as well as English-speaking controls), we examine how this perception-production process is influenced by language. In the first study, native Japanese speakers produced an English word with formant-shifted feedback. In the second experiment, native Japanese speakers produced a Japanese syllable with altered feedback and produced an

English word that contained a similar vowel with normal feedback. The results were compared with data from English controls and suggest that the compensatory behavior is not language dependent.

4pSW29. Linguistic experience's influence on foreign accent detection in short, slightly accented speech. Hanyong Park (Speech Res. Lab., Indiana Univ., 1101 E. 10th St., Bloomington, IN 47405, hanyongpark@indiana.edu)

This study examined how a listener's linguistic experience affects the detection of foreign accent in short, slightly accented speech. Two factors were considered: Age of arrival in a second-language (L2) speaking country (AOA), and length of residence in that country (LOR). Two listener groups with the same native language (L1: Korean) but with different AOA and LOR judged the nativeness of short stimuli produced by 4 Korean-English bilinguals and 2 native speakers of American English. The stimulus length ranged from the vowel /a/ to monosyllabic and disyllabic English words. To investigate the listeners' sensitivity patterns to different linguistic structure, the monosyllabic corpus included stimuli having both natural (i.e., CV) and unnatural syllable structures (i.e., CCV, CVC, and CCVC) as well as various segments in terms of Korean phonotactics. Results show that: (1) more experienced with L2 led to higher sensitivity to a foreign accent; (2) AOA affects foreign accent detection more than LOR; and (3) the non-native listeners were not sensitive to different linguistic structures, except in different CCV types by the more experienced listeners. The results suggest that a listener's sensitivity to a foreign accent develops up to a certain degree, according to L2 experience.

4pSW30. Asymmetries in the mismatch negativity response to vowels by French, English, and bilingual adults: Evidence for a language-universal bias. Linda Polka, Monika Molnar, Shari Baum (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine Ave. W., Montreal, QC Canada, linda.polka@McGill.ca), Lucie Menard (Universite du Quebec à Montréal, QC Canada), and Karsten Steinhauer (McGill Univ., Montreal, QC, Canada)

In infants, discrimination of a vowel change presented in one direction is often significantly better compared to the same change presented in the reverse direction. These directional asymmetries reveal a language-universal bias favoring vowels with extreme articulatory-acoustic properties (peripheral in F1/F2 vowel space). In adults, asymmetries are observed for non-native but not for native vowel contrasts. To examine neurophysiological correlates of these asymmetries we recorded MMN responses from monolingual (English, French) and simultaneous bilingual (English/French) adults using an oddball paradigm with four vowels: English [u], French [u], French [y], and an acoustically-distinct [y]. All vowels were tested in four conditions with each vowel serving as deviant and standard. Within each vowel pair, MMN responses were larger and earlier when the deviant vowel was more peripheral. This pattern was consistent across the language groups and was observed for within-category (within [u]; within [y]) and for cross-category vowel pairs ([u] versus [y]). Findings indicate that a bias favoring peripheral vowels is retained in the neural pre-attentive processing of vowels in adults. As in infant behavioral findings, this bias is independent of the functional status of the vowel pair. Implications for the natural referent vowel model [Polka and Bohn (2003)] will be discussed.

4pSW31. The effect of accent on toddlers' story comprehension and word recognition. Brittan A. Barker and Lindsay E. Meyer (Dept. of COMD, Louisiana State Univ., 63 Hatcher Hall, Baton Rouge, LA 70803)

The current study sought to enrich the developmental research and contribute to the ever-growing knowledge about talker-specific spoken language processing by examining the affect of accent on toddlers' ability to comprehend a story. A total of 24 children aged 30–42 months participated in the first study employing a between-subjects design. Talker accent (native versus non-native) served as the independent variable and story comprehension accuracy served as the dependent variable. It was predicted that the children listening to the story in their native accent would perform significantly better on the comprehension task than those listening to the non-native accent. The hypothesis was not supported; there was no significant difference between the listeners' performance [$t(22) = .169, p > .05$]. An ongoing, follow-up study was conducted to further explore the surprising results. The affect of accent on word recognition skills in these same children is currently being tested. The complete data set of the second study will be presented. It is hypothesized that the toddlers will have significantly more difficulty recog-

nizing the words spoken in a non-native accent. These results would echo previous talker-specific word recognition results [Ryalls and Pisoni (1997)] and suggest that talker-specific information may affect low-level speech perception but not higher-level language comprehension.

4pSW32. Cross-linguistic variation in language similarity classification. MaryAnn Walter (Dept. of Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208, m-walter@northwestern.edu)

This study aims at identifying factors that make language sound structures seem more or less similar to English, and how those similarity judgments change according to the listener's native language. Listeners from four different native language groups (English, Mandarin Chinese, Korean, and Turkish) sorted a group of 17 genetically and geographically diverse languages in terms of their sound-based distance from English. Placements of individual languages were analyzed, as well as groupings of similarly ranked languages and correlations among overall ranking structures of the different groups. Bilingual listeners exhibit more variability in their rankings than monolingual English speakers, rank their own language as less similar to English than other groups do, and rank languages of neighboring groups the least similar to it of all. Ranking correlations between language groups are significant, varying somewhat in magnitude depending on geographical proximity and typological/genetic relatedness of the listener group languages. This reflects the presence of consistent groupings within ranking structures for all language groups, which depend on sound-based factors such as the presence of perceptually salient speech sounds. These results enable predictions about relative intelligibility among international English users, native and non-native. [Work supported by a Mellon Postdoctoral Fellowship awarded by the Woodrow Wilson Foundation.]

4pSW33. Acoustic similarities between front rounded and back unrounded vowels as evidenced by French /ø/ and /u/ produced by Japanese-speaking learners. Takeki Kamiyama (Laboratoire de Phonétique et Phonologie (UMR 7018) CNRS/Sorbonne Nouvelle, 19, rue des Bernardins, 75005 Paris, France, Takeki.Kamiyama@univ-paris3.fr)

French high back rounded /u/ is characterized by a concentration of energy in the low frequency zone (< 1000 Hz) due to the grouping of the first two formants, and midhigh front rounded /ø/ by a balanced distribution of formants, with F2 located around 1500 Hz. Japanese-speaking learners of French (JSL), who have difficulty differentiating /ø/ and /u/ both in perception and production, tend to produce, for both, Japanese-like /u/, which is fronted and less rounded, with a formant structure similar to that of French /ø/. Our perception experiment using 18 tokens each of /u y ø/ produced by five JSL shows that the 16 native speakers of French (NF) tested perceived mainly /ø/ when they heard those stimuli intended as /u/ by JSL but produced with a high F2 between 1100 and 1600 Hz, with a mean goodness rating of 2–4.5 out of 5 for /ø/. Another perception test conducted with stimuli synthesized using Maeda's articulatory synthesis (VTCalcs) shows that the 16 NF examined tended to identify vowels synthesized as front rounded vowels and back unrounded vowels both as front unrounded ones (/ø/ and /æ/, respectively), which indicates the acoustic similarities of both types of vowels.

4pSW34. Relation of perception training to production of codas in English as a second language. Teresa Lopez-Soto (Dept. of English Lang., Univ. of Seville, Palos de la Frontera, Sevilla 41004, Spain and Indiana Univ., Bloomington, IN 47405) and Diane Kewley-Port (Indiana Univ., Bloomington, IN 47405)

A preliminary study has been conducted to discover whether moderate amounts of speech perception training improve accurate production even though production is not trained. The study recruited one group of eight Spanish adults who had resided less than 10 years in the USA. A set of 13 word-final English consonants was selected for training from a SPATS software module. On days 1 and 5, the group participated in both perception and production tasks with the 13 codas (pre- and post-tests). On days 2–4, the group trained with feedback for 1 h mostly with VC syllables and occasionally with sentences (speech recorded from multiple talkers). Results show: (1) with 3 h training, Spanish listeners' perception improved significantly across the 13 codas, with greater improvement on consonant clusters than on singletons; (2) for consonants not accurately produced in the pretest, many substantially improved after only perception training; (3) several consonants

with large gains in perception also showed the large improvements in production. The results from this study suggest that training only with perception can improve speech production. Experience with this protocol lays the groundwork for a series of studies to examine how perception and production are linked in learning a new language.

4pSW35. Confusion direction differences in second language production and perception. Yen-Chen Hao and Kenneth de Jong (Dept. of Linguist., Indiana Univ., 322 Memorial Hall, 1021 E. 3rd St., Bloomington, IN 47405)

This study examines differences between confusions found in productions and perceptions of learners of English. Twenty Korean EFL learners engaged in three tasks involving obstruents placed in different prosodic positions: (a) identification of native English productions, (b) reading from orthographic prompts, and (c) mimicry of native English productions. Recordings of reading and mimicry were presented to 50 native English listeners for identification. This paper compares patterns of errors found for 10 intervocalic obstruents before and after a stress, since previous studies showed that Korean does not exhibit stress-induced differences in consonant allophones. Similarity estimates using Luce's similarity choice model were regressed across the two intervocalic positions. We found robust correlations despite allophonic differences in English, suggesting a component of L1 transfer in all three tasks. Examining bias parameters, however, revealed systematic differences in the direction of the resulting errors, which is task-dependent. Identification and mimicry tended to underestimate allophonic shifts due to stress, and so to create more voiceless to voiced errors in post-stress environments. Reading productions exhibited error directions in exactly the opposite directions, suggesting Korean learners produced the stress but not the corresponding allophonic variations. These patterns indicate very different error outcomes in production and perception.

4pSW36. Maintenance of /e-ε/ in word-final position as a phonemic and morphemic contrast in Canadian French. Franzo Law, II and Winifred Strange (Program in Speech-Lang.-Hearing Sci., CUNY Graduate Ctr., 365 5th Ave., New York, NY 10016, flaw@gc.cuny.edu)

Many dialects of French have merged /e-ε/ to /e/ in word-final context. The present study investigated the stability of this contrast in Canadian French. Productions of four Canadian French-dominant speakers (two monolingual, two bilingual) were recorded and analyzed. Real (e.g., "thé") and nonsense (e.g., "gispaïs") words ending with /e/ or /ε/ were used, as well as real and nonsense verbs distinguished morphosyntactically by the same vowel contrast (e.g., first person singular future "parlerai" versus first person singular conditional "parlerais"), all embedded in carrier sentences. Results showed that participants maintained spectral and duration distinction for the vowel contrast when preceded by labial, coronal, and back stops in real and nonsense words. All participants showed varying degrees of coarticulation for /e/ when preceded by /t/; /ε/ was spectrally lower than in other contexts. Three of the four participants maintained a stable distinction in morphosyntactic context: monolinguals exhibited the best retention of the distinction, whereas the bilinguals had partially or completely overlapping distributions, merging /e-ε/ to /e/, not /ε/. These patterns suggest difficulties in perception of this contrast for English late L2 learners of French, due to the fact that this contrast is not phonotactically possible in English. [Work supported by NIH F31DC008075.]

4pSW37. On prosodic variation and pronunciation aptitude: A case study of German speakers of English. Volha Anufryk and Grzegorz Dogil (Azenbergstr. 12, 70174 Stuttgart, Germany)

The present study investigates prosodic variation, as realized by L1 German speakers of varying pronunciation aptitude (below-average, average, and above-average) in comparison with native speakers of English. The results demonstrate the frequency of distribution of various pitch contours in read and spontaneous speech samples, English and German, on both the phonological (ToBI pitch accents and boundary tones) and phonetic levels. The latter is analyzed using a parametric intonation model (Moehler 2001). On the phonological level the rising contours, i.e., boundary tones and pitch accents, have a wider distribution in German as well as English productions of the German subjects as compared to the L1 English speakers. In German the percentages in all the aptitude groups are almost equal covering about 50% of all contours. Their English language realizations, on the other hand, are marked by a scarcer distribution of the rising pattern, with the highest

percentages in the below-average and average groups and lowest by the above-average speakers. Consequently, there is a clearcut tendency for the native pattern in the German speakers' samples. Nonetheless there is evidence for accommodation to the English variation pattern, which is strongest in the above-average aptitude group.

4pSW38. Native and non-native vocal imitation by vocalists and nonvocalists. Terry L. Gottfried and LaDonna Hayden (Dept. of Psychol., Lawrence Univ., P.O. Box 599, Appleton, WI 54912, gottfrit@lawrence.edu)

Several studies [e.g., Slevc and Miyaki, Psychol. Sci. 17, 675–681 (2006)] have demonstrated that musical ability predicts phonological accuracy in second language learning. The present study examined whether vocal musicians' diction training affected their immediate imitation of familiar (English nonwords) and unfamiliar (Mandarin) syllables. Participants were 19 nonmusic majors (recruited from introductory linguistics) and 19 vocal music majors (who had completed three of four vocal diction classes). All participants heard a list of eight English nonwords and eight Mandarin words from a recording and repeated each word or nonword to the best of their ability. These recordings were rated (by two bilingual judges) on the overall accuracy of word repetition and vowel and consonant accuracy. Participants, regardless of group, repeated English nonwords more accurately than Mandarin words. Vocalists repeated both English nonwords and Mandarin words more accurately than the nonvocalists. This could indicate that vocalists' experience in diction training improved their ability to repeat any speech sound. Acoustical analyses of these vocal imitations are currently being conducted, and future studies will examine whether musicians without diction training are better than nonmusicians in vocal imitation.

4pSW39. Gestural drift in Spanish-English speakers. Stephen Tobin (Dept. of Psych., Univ. of Connecticut, 406 Babbidge Rd., Unit 1020, Storrs, CT 06269-1020, stephen.tobin@uconn.edu)

Voice onset times (VOTs) and second vowel formants (F2s) of native speakers of Spanish were examined to extend findings on perceptually guided changes in speech production [M. L. Sancier and C. A. Fowler, J. Phon. 25, 421–426 (1995)]. Participants were recorded producing English and Spanish sentences after stays in the U.S. and in Spanish-speaking countries. Participants' English and Spanish VOTs were both significantly shorter after a stay in a Spanish-speaking country than after a stay in the U.S., supporting previous findings. While S1's F2s were significantly lower following exposure to Spanish, as expected based on earlier results [A.R. Bradlow, J. Acoust. Soc. Am. 97(3), 1916–1924 (1995)], S2's F2s increased. We attribute this effect to gender, pending further data acquisition. We propose that not only inter-gestural timing but also the gestural targets themselves are affected by ambient language. The systematic effect of language exposure on F2 is consistent with language-specific articulatory settings [B. Honikman (1964); Gick et al., Phonetica 61, 220–233 (2004)].

4pSW40. Vowel duration in Mexican heritage English. Kenneth Konopka and Janet Pierrehumbert (Dept. of Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208, k-konopka@northwestern.edu)

Mexican heritage English (MHE) speakers in Chicago comprise a unique population of native English speakers; their speech is often regarded by non-ethnic (NE) speakers as Spanish-accented English in spite of the fact that they may have low or no proficiency in Spanish. Vowels were analyzed from recordings of wordlist readings by MHE, NE, and adult Mexican ESL speakers (L2E) from a single Chicago community ($N=15F, 10F, 12F$, respectively). The analyses indicate that MHE vowel durations correspond to NE durations for the dimoraic (long) vowels, but correspond to the Spanish-accented vowels of L2E for the monomoraic (short) vowels. Thus the range of duration between short and long vowels is compressed for MHE vowels, reflecting the NE norm for long vowels, and the Spanish-influenced minimum for short vowels. Sociophoneticians often rely on vowel formant frequency plots to provide a quantitative measure of dialect identity that may correlate with a variety of regional and social factors. The current study provides evidence that vowel duration is a phonological contrast in the MHE dialect that reflects the influence of language contact on vowel properties that are not captured by vowel plots alone.

4pSW41. The perception of intermediate phonological relationships. Kathleen Currie Hall (Dept. of Linguist., Ohio State U., 1712 Neil Ave, Columbus, OH 43210, kchall@ling.osu.edu)

It has been shown that pairs of segments that are allophonic in a language are perceived as being more similar than pairs that are contrastive in a language [Boomershine et al., (2008)]. There is also evidence that neutralized contrasts in a language are perceived as more similar than non-neutralized contrasts [Hume and Johnson (2003)]. Third, there is evidence that phonological relationships should be defined along a continuum of predictability, rather than as a categorical distinction between “allophonic” and “contrastive” [Hall (2008)]. In combination, these facts predict that pairs of segments that fall along a cline of predictability of distribution should also

fall along a cline of perceived similarity. This paper presents results of a perception experiment that tests this prediction by examining the perceived similarity of four pairs of sounds in German: (1) [t]-[tʃ], which is almost fully contrastive (unpredictably distributed); (2) [t]-[d] and (3) [s]-[ʃ], which are each partially contrastive (partially predictably distributed); and (4) [x]-[ç], which is almost fully allophonic (completely predictably distributed). If the notion of a cline of predictability is correct, these four pairs will align themselves along a cline of similarity with [t]-[tʃ] being rated as the most distinct and [x]-[ç] as the most similar.

FRIDAY MORNING, 22 MAY 2009

SKY BRIDGE AUDITORIUM, 8:45 TO 10:10 A.M.

Session 5aSWa

Speech Workshop: New Perspectives on Models of Cross-Language Speech Perception

Ann R. Bradlow, Chair

Dept. of Linguistics, Northwestern Univ., Evanston, IL 60208

Chair's Introduction—8:45

Invited Papers

9:00

5aSWa1. Accounting for the accented perception of vowels: Universal preferences and language-specific biases. Ocke-Schwen Bohn (English Dept., Aarhus Univ., J.-C.-Skous Vej 5, DK-8000 Aarhus C, Denmark) and Linda Polka (School of Commun. Sci. and Disord., Montreal, PQ, H3G1A8, Canada)

Strange things happen in cross-language and second-language vowel perception: Naïve non-native listeners have been reported to rely on acoustic properties which are nonfunctional in their L1 and dysfunctional for the perception of non-native vowels; naïve non-native listeners' perception is guided by a preference for vowels that are peripheral in the articulatory/acoustic vowel space; and, in general, naïve non-native listeners' perception is not well predicted by comparative analyses of vowels of the native and the non-native language. This presentation reviews the accented perception of vowels by focusing on two forces which shape non-native vowel perception: universal perceptual preferences which non-native listeners (and infants) bring to the task of vowel perception, and perceptual biases which non-native listeners transfer from their native to the non-native language. Strange and her colleagues have shown that these biases cannot be predicted from acoustic comparisons; rather, they have to be examined directly through assessments of the perceived cross-language similarity of vowels. This presentation addresses several of the still unresolved questions regarding the design and the interpretation of perceptual assimilation tasks used to account for the accented perception of vowels. [Work supported by Danish Research Council for the Humanities, Canadian Natural Sciences and Engineering Research Council.]

9:35

5aSWa2. Articulating the Perceptual Assimilation Model (PAM): Perceptual assimilation in relation to articulatory organs and their constriction gestures. Catherine T. Best (MARCS Auditory Labs, Univ. Western Sydney, Locked Bag 1797, Penrith NSW 1797, Australia, and Haskins Labs, 300 George St., New Haven, CT 06511, c.best@uws.edu.au), Louis Goldstein (Univ. Southern Calif., Los Angeles, CA 90089-1693), Michael D. Tyler (Univ. Western Sydney, Penrith NSW 1797, Australia), and Hosung Nam (Haskins Labs, New Haven, CT 06511)

A core premise of the Perceptual Assimilation Model of non-native speech perception (PAM) [Best (1995); Best & Tyler (2007)] is that adults perceive unfamiliar non-native phones in terms of articulatory similarities/dissimilarities to native phonemes and contrasts. The implied attunement to native speech emerges early: As infants begin to discern the articulatory organization of native speech, language-specific effects in non-native speech perception appear (≈6–10 months). Given that non-native phones, by definition, deviate phonetically from native ones, how can we characterize articulatory similarity in concrete, testable ways? The Articulatory Organ Hypothesis (AO) [Studdert-Kennedy & Goldstein (2003); Goldstein & Fowler (2003)] offers a possible approach, positing that infants decompose the oral-facial system into distinct articulatory organs (e.g., lips, tongue tip, tongue dorsum) and are sensitive to their actions in producing vocal tract constrictions. Thus, between-organ contrasts should be easily perceived/learned by infants and adults, whereas detection of within-organ contrasts must become attuned to the distribution of differing constriction locations/types by that organ in input speech. We discuss articulatory, attunement modeling, and perceptual evidence consistent with these predictions, and present a revised version of PAM that incorporates the AO Hypothesis and related principles of articulatory phonology [Browman & Goldstein (1991)]. [Work supported by NIH.]