Session 2pSC

Speech Communication: Speech Production (Poster Session)

Rajka Smiljanic, Chair
Northwestern Univ., Dept. of Linguistics, 2016 Sheridan Rd., Evanston, IL 60208

Contributed Papers

All posters will be on display from 1:30 p.m. to 4:30 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:30 p.m. to 3:00 p.m. and contributors of even-numbered papers will be at their posters from 3:00 p.m. to 4:30 p.m.

2pSC1. Finite-element simulations of self-oscillation mechanisms in human phonation. Douglas Cook, Luc Mongeau, and Paul Smith (Purdue Univ., 585 Purdue Mall, West Lafayette, IN 47907)

The fluid-structure interactions which occur during human phonation were investigated using a commercially available finite-element program. Both lumped mass and continuous structural models were coupled with viscous fluid flow. The simulated behavior was analyzed. The addition of viscous flow was not sufficient to induce self-oscillation of a single degree-of-freedom lumped mass model but did contribute to the steady vibration of the continuous model. This observation reinforces the conclusion that structural complexity is an essential mechanism of self-oscillation. The influence of false folds on the oscillation patterns of the continuous model was also examined.

2pSC2. Characteristics of phonation in women who produce voiced versus voiceless /h/. Laura L. Koenig (Haskins Labs., 300 George St., New Haven, CT 06511 & Long Island Univ., Brooklyn, NY 11201)

Production of the consonant /h/ typically involves a vocal-fold abduction gesture. Although /h/ is traditionally described as voiceless in English, voiced allophones also occur. Our recent work has described the phonatory behavior of women who produce a mixture of voiced and voiceless /h/; our results suggested speaker-specific patterns of the factors that lead to devoicing. This study characterizes the behavior of normal women who produce exclusively voiced /h/ and compares it to those who produced both voiced and voiceless /h/. Measures of baseline (dc) airflow are used as an index of abduction degree; peak intraoral pressures during adjacent /p/ closures provide a measure of subglottal pressure; and pulse-by-pulse measures are made of $f0$ and pulse amplitude (ac flow) around the /h/ abduction gesture. Multidimensional analyses are carried out within speakers to determine how these vocal parameters are related. We expect that women who produce only voiced /h/ will differ from those who devoice some /h/ productions in the parameters that facilitate phonation, such as higher subglottal pressures, smaller abduction degrees, and reduced vocal-fold tension as reflected in $f0$ measures. Based on our previous findings, however, we also expect individual variation in the phonatory settings that produce voiced /h/. [Work supported by NIH.]

2pSC3. Contextual effects on the continuancy of /ð/. Sherry Zhao (MIT, 77 Massachusetts Ave., Rm. 36-545, Cambridge, MA 02139)

This research examined the segmental-contextual effects on the continuancy of the voiced dental /ð/ in American-English and whether certain acoustic attributes were preserved despite possible modification. Word-initial /ð/ cases, extracted from continuous speech of 146 speakers from the TIMIT database, were frequently (averaging 65%) stoplike when it was in utterance-initial position or when its preceding phoneme was voiceless and/or [-continuant]. This stoplike modification occurred less frequently (averaging 39%) when /ð/ was preceded by a voiced fricative and rarely (averaging 13%) when preceded by a vowel or liquid consonant. A comparison of stoplike /ð/ and /d/ cases under similar contexts showed that the burst peak location, burst spectrum shape, and F2 at vowel onset averages were all statistically different between the two groups. In addition, the acoustic data suggested that the dental place of articulation was preserved for the modified /ð/. Preliminary automatic classification experiments involving salient acoustic attributes seemed to indicate that F2 at vowel onset may be a reliable cue for the dental place of articulation in /ð/. [Work supported by NIH DC02978 and T32DC00038.]

2pSC4. Modeling the influence of acoustic loading on laryngeal self-sustained oscillations. Matías Zañartu, Luc Mongeau, and George R. Wodicka (Ray W. Herrick Labs. Purdue Univ., West Lafayette, IN 47907)

Three-way interactions between sound waves in the subglottal and supraglottal tracts, the vibrations of the vocal folds, and laryngeal flow were investigated. Nonlinear coupling was investigated analytically, with a focus on the effects of a realistic acoustic loading on the self-sustained oscillation of the vocal folds. MRI-based shapes for the supraglottal tract [Story *et al.*, J. Acoust. Soc. Am. **100**, 537–554 (1996)] and an equivalent acoustical model for the subglottal tract [Harper *et al.*, IEEE Trans. Biomed. Eng. **48**, 543–550 (2001)] were considered. The model was based on the wave reflection analog technique, along with a novel single degree of freedom model of the vocal folds to describe the source. The effects of a mucosal wave on the vocal folds were incorporated through a time-varying discharge coefficient and a Bernoulli-type flow model. The relative importance of acoustic loading and time-varying flow resistance for fluid-structure energy transfer was compared for various configurations. [Work supported by NIH Grant RO1-DC05788.]

2pSC5. Patterns of interarticulator coordination within the syllable. Patrizia Bonaventura and Joseph Sand (Dept. of Commun. Sci., Case Western Reserve Univ., 11206 Euclid Ave, Cleveland, OH 44106, patrizia.bonaventura@case.edu)

Interarticulatory coordination patterns were observed in movements of the crucial articulator for the production of place in selected syllables containing labiodental and alveolar consonants. The coordination of the lower lip with upper lip and tongue tip, and the coordination of the tongue tip with tongue blade and lower lip have been observed. The goals of the study were (a) to verify whether stable patterns of speed with respect to excursion of movement [Fujimura, O., "Temporal organization of articulatory movements as a multidimensional phrase structure," Phonetica, **38**, 66–83 (1981a); "Icebergs revised," ASA meeting, 2aSC (1996b)], were

found in concurrent articulators to the crucial one for place, under different prosodic conditions; (b) whether such stable portions would vary proportionally to the increase in amplitude of the vertical jaw movement, selected as a measure of the magnitude of the syllable; (c) to observe relative timing at the iceberg thresholds in the selected consonantal movements with respect to the center of the syllable. Analyses were performed on articulatory recordings from cinefluorographic data, based on read dialogues [Erickson *et al.*, "Articulatory correlates of prosodic control: Emotion and emphasis," Lang. Speech, **41**, 399–417 (1998)]. Iceberg patterns were found in the concurrent articulator movements, showing a linear dependence of slope on the total excursion of the demisyllabic movement. Timing of concurrent articulator movements with respect to the center of the syllable is shown to be relatively stable.

2pSC6. Modulation of frication noise in a dynamic mechanical model of the larynx and vocal tract. Anna Barney (ISVR, Univ. of Southampton, Southampton, SO17 1BJ, UK, ab3@soton.ac.uk) and Philip J. B. Jackson (Univ. of Surrey, Guildford, GU2 7XH, UK)

A dynamic mechanical model of the larynx and vocal tract was used to investigate the aerodynamics and acoustics of voiced fricatives. The model had typical dimensions for a human adult male. Glottal excitation arose from driven shutters representing the vocal folds; sinusoidal and square-wave driver functions were investigated. Glottal amplitude, fundamental frequency ($f0$), and volume velocity were independently controlled. Frication noise was produced by an orifice plate with a sharp-edged obstacle on its downstream side. Two different positions of constriction within the vocal tract, each coupled with a different constriction-obstacle distance, were investigated. Vocal-tract wall pressure was measured at three locations and a microphone at the vocal-tract exit measured radiated sound. The sound radiated from the open end of the vocal tract was found to have both harmonic and noise components. The noise was modulated by the $f0$ of the glottal vibration. The depth of modulation was independent of the shape of the glottal excitation. It showed an overall decreasing trend with increasing $f0$ in the range 20 to 150 Hz, but with a local maximum of modulation depth around 80 Hz. The noise bursts had maxima coincident with the middle of the glottal open phase.

2pSC7. Acoustic characteristics of clearly produced fricatives. Kazumi Maniwa, Allard Jongman (Linguist. Dept., Univ. of Kansas, 1541 Lilac Ln., Blake Hall, Rm. 427, Lawrence, KS 66044-3177), and Travis Wade (Posit Science Corp., San Francisco, CA 94104)

Research suggests that speakers can adopt a speaking style that allows them to be understood more easily when confronted with difficult communication situations, but few studies have examined the acoustic properties of clearly produced consonants in detail. This study attempts to characterize the type and magnitude of adaptations in the clear production of English fricatives in a carefully controlled range of communication situations. Ten female and ten male talkers produced nonsense syllables containing the eight English fricatives in VCV contexts, both in a conversational style and in a clear style (elicited by means of feedback consisting of simulated recognition errors received from an interactive computer program). Acoustic measurements were taken for spectral, amplitudinal, and temporal properties known to influence fricative recognition. Results illustrate that (1) there were some consistent overall clear speech effects, several of which (consonant duration, spectral peak location, spectral moments) were consistent with previous findings and a few (notably consonant-to-vowel intensity ratio) which were not; (2) "contrastive" differences related to eliciting prompts were observed in a few key comparisons; and (3) talkers differed widely in the types and magnitude of acoustic modifications.

2pSC8. Articulation of fricatives: Evidence from X-ray microbeam data. Christine H. Shadle, Khalil Iskarous, and Michael I. Proctor (Haskins Labs., 300 George St., New Haven, CT 06511)

Articulatory changes in /s/ over time give rise to the onset and offset of turbulence noise, but the small tolerances involved make it difficult to separately consider the effects of movement, phonetic context, and subject variation. Aerodynamic estimates of the constriction area have been shown to demonstrate a marked asymmetry, with the /s/ onset more rapid than offset [C. Scully, Speech Commun. **11**, 1992]. Previous articulatory studies led us to expect an invariant tongue tip position, with jaw raising dependent somewhat on the following vowel. In order to investigate the timing of the constriction formation in fricatives, X-ray microbeam data [Westbury *et al.*, 1994] for citation forms of /sVd/ for 15 vowels were studied. In fact, the tongue marker patterns during [s] vary both by following vowel and subject; subjects seem to have different apical-laminal strategies. The jaw reaches maximum height during [s] for each subject studied; that height does not appreciably vary with vowel. This indicates that jaw height is important specifically for /s/ production. One possibility is that the constriction is formed by the tongue, and turbulence noise is then maximized by the lower teeth being brought into position. [Work supported by NIH NIDCD R01 DC 006705.]

2pSC9. A range of intonation patterns produced in an elicitation task. Alejna Brugos, Jonathan Barnes (Program in Appl. Linguist., Boston Univ., 96 Cummington St. #246, Boston, MA 02215, abrugos@bu.edu), Stefanie Shattuck-Hufnagel (MIT, Cambridge, MA 02139), and Nanette Veilleux (Simmons College, Boston, MA 02115)

A production task was designed to elicit a specific intonation contour, characterized by an exaggerated rise-fall-rise in $f0$ ($L+H^*L-H\%$ in ToBI terms), conveying an attitude of "dismayed surprise." Speakers were given models, a short practice period, and a preceding context for each utterance. A targetlike rise-fall-rise contour (including $H^*L-H\%$ variants) was successfully elicited for each of the contexts, but in less than half the utterances: Of 1134 utterances, 631 (56%) were judged by a panel of 3 experienced ToBI labelers as different from the intended contour, e.g., $L^*H-H\%$. Of these, 261 were produced with contours perceptually similar to the target, with a sustained flat or minimally rising $f0$ after the fall from the pitch accent, or with the rise occurring early in the next phrase. Alignment characteristics and individual speaker data may determine whether these are one contour with a continuum of realizations, or separate contours. These results (a) suggest that speakers have more than one way of signaling a given attitude and resist being limited to using only one, and (b) support the need for careful postexperiment transcription to characterize which contours subjects have actually produced. [Work supported by NSF 0345627 and NIH DC00075.]

2pSC10. Factors affecting phonetic variation of American English /k/. Stefanie Shattuck-Hufnagel (Speech Group, Res. Lab. of Electron., MIT, Cambridge, MA 02139) and Nanette Veilleux (Simmons College, Boston, MA 02115)

Word-final consonant reduction is commonly observed in American English, but variation in word-initial consonants is less thoroughly studied. Several considerations suggest that initial consonants might resist reduction; a more nuanced hypothesis is that initial consonants resist reduction in prosodically strong positions (i.e., at phrase edges and pitch accents) but may be reduced in weaker positions (e.g., phrase medially, in nonaccented syllables), particularly in un-self-conscious speech. A corpus of 16 spontaneous dialogues (8 speakers), prosodically labeled using the ToBI transcription system, provides a testbed for this view: the speakers were pairs of close friends engaged in an absorbing task which distracted them from paying close attention to their speech. Preliminary inspection of initial /k/'s from 3 speakers revealed a range of pronunciations (as described in earlier work), from normal prevocalic /k/ (with a silent closure period followed by release noise), to "leaky" /k/'s with some noise during the closure period and higher-amplitude release noise, to fricated /k/'s with

constant-amplitude noise throughout, and offered some support for reduction in weak prosodic contexts. Further analysis will compare prosodic context with other factors governing phonetic variation (e.g., word frequency); implications for speech recognition and production will be discussed. [Work supported by NIH DC00075 and DC02978.]

2pSC11. Errors and strategy shifts in speech production indicate multiple levels of representation. Bryan Gick (Dept. of Linguist., Univ. of BC, E270-1866 Main Mall, Vancouver, BC V6T1Z1 Canada)

An earlier pilot study [Gick, J. Acoust. Soc. Am. **112**, 2416 (2002)] observed that speakers producing challenging sequences of articulatory movements exhibit categorical shifts in production strategies. For example, a speaker may initiate all two-flap sequences with an upward tongue movement, then shift to a downward movement, then shift again. The present study attempts to determine how these strategy shifts can inform phonologists as to levels of representation. Ten subjects produced sequences of flap allophones of English /t/ and /d/ in an experimental paradigm similar to the pilot study. Results show strategy shifts in the speech of 7/10 subjects, such that: (1) shifts are more likely to occur following speech errors; (2) errors are more likely to occur in sequences with more flaps and more conflicts; (3) errors in response to articulatory conflicts occur at multiple levels of representation (segmental, lexical, etc.); (4) strategy shifts occur at multiple levels of representation (subsegmental, segmental, whole word); (5) word frequency does not affect strategy shifts. These results suggest that speech processing is active in parallel at multiple levels associated with phonological representation. [Work supported by NSERC.]

2pSC12. 3-D acoustic analysis based on magnetic resonance imaging of a "retroflex" and "bunched" American English rhotic sound. Xinhui Zhou, Carol Espy-Wilson (Dept. of Elec. and Comput. Eng., Univ. of Maryland, 2205 AV Williams Bldg., College Park, MD 20742, espy@glue.umd.edu), Mark Tiede (Haskins Labs, New Haven, CT 06511), and Suzanne Boyce (Univ. of Cincinnati, Cincinnati, OH 45202)

A variety of articulatory configurations can produce an American English /r/. Common across all configurations is a large front-cavity volume that results directly or indirectly in the characteristically low third formant. Earlier work [Zhang *et al.*, ICASSP2005 1, pp. 893–896] using 2-D finite-element modeling of the front cavity with stylized bunched and retroflexed shapes showed that, even at 2000 Hz, cross modes are generated in the front cavity. Thus, the front cavity should be modeled as a single tube with a side branch, regardless of whether there is space underneath the tongue. This paper extends the earlier work by (a) considering the full 3-D vocal tract geometry reconstructed from magnetic resonance images, (b) performing 3-D finite-element analysis, and (c) using contrasting tongue shapes from similar-sized subjects. Harmonic analysis showed the frequency response is comparable to the spectrum of the acoustic signal, suggesting good agreement between the actual and reconstructed vocal tracts. As before, wave propagation in the front cavity is not planar. 3-D vocal tract analysis shows that the cross modes have 3-D properties. Application of the 3-D wave propagation property to area function extraction and simple-tube modeling of the vocal tract will be discussed. [Research supported by NIH.]

2pSC13. Syllable structure effects on velum-oral coordination evaluated with real-time MRI. Stephen Tobin, Dani Byrd (USC Linguist., 3601 Watt Way, GFS 301, Los Angeles, CA 90089-1693, dbyrd@usc.edu), Erik Bresch, and Shrikanth Narayanan (USC Los Angeles, CA 90089)

Seminal work by Krakow (PhD. Diss., Yale University, 1989, Phonetics and Phonology **5**, 87–116, 1993) has indicated that intrasegmental coordination of articulatory gestures varies systematically as a function of syllable position. Specifically, onset and coda nasals display different lag times between the velum and oral gestural components. In onset nasals, velum lowering is achieved roughly synchronously with oral constriction, while in codas, velum lowering precedes oral constriction. However, Krakow's study was limited by its reliance on indirect measures of velum movement. Nor did her work examine nasals other than [m] or syllable organizations other than singleton onsets and codas. Our study employs real-time magnetic resonance imaging [Narayanan *et al.*, J. Acoust. Soc. Am. **115**, 1771–1776 (2004); Mathiak *et al.*, Int. J. of Lang. & Comm. Disorders **35**, 419–425 (2000)] to acquire direct information about the key vocal tract variables of velum aperture and oral constriction, thereby allowing the calculation of relative gestural timing. We investigate the effects of syllabification on interarticulatory timing in [n], and, in addition to examining codas and onsets, further consider juncture geminates spanning word boundaries. Preliminary real-time MRI results from one of three anticipated subjects do find the predicted timing difference for [n]. Further, we find that the relative timing in juncture geminates mirrors that of coda [n] but presents a velum aperture intermediate between coda and onset [n]. [Work supported by NIH.]

2pSC14. Adaptation of speech production to palatal perturbation: Kinematic analysis using electromagnetometry. Douglas M. Shiller (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine Ave. West, Montreal, QC, Canada H3G 1A8, doug.shiller@mail.mcgill.ca), Mark K. Tiede (Haskins Labs., New Haven, CT), and Vincent L. Gracco (McGill Univ., Montreal, QC, Canada)

The extent to which adults are able to flexibly adapt their speech production to altered sensory-motor conditions has considerable implications for oral-motor rehabilitation, yet it remains poorly understood. Previous studies have explored the effects of different physical and perceptual perturbations on speech production, yielding mixed results with respect to the timecourse and extent of adaptation. A major limitation of these studies has been their reliance on acoustical measures as the sole index of motor adaptation. The difficulty with such an approach is the potentially complex relationship between acoustic changes and their underlying articulatory patterns. In the present study, speech adaptation to a palatal perturbation (5-mm thickened alveolar ridge) was investigated by directly measuring articulatory kinematics using electromagnetometry. In contrast to previous claims, subjects were found to produce a significant change in speech-related tongue movement immediately upon insertion and removal of the artificial palate. In addition, subjects gradually fine-tuned their tongue position with continued practice. The observed patterns are consistent with the idea of two independent adaptation processes: a rapid adaptation process that is driven by the mere presence of the false palate and a slower adaptation process driven by the continued presence of the palatal perturbation.

2pSC15. Production of American English word-initial stops in the vicinity of a voiceless stop. Elina Eydlin (Dept. of Speech-Lang. Pathol. and Audiol., New York Univ., 719 Broadway, Ste. 200, New York, NY 10003) and Nassima Abdelli-Beruh (New York Univ., New York, NY 10003)

This study examined how native-American English speakers produced the stop voicing distinction when word-initial stops were positioned before a voiceless stop (/t/) in the sentence repeat CVC ten times. Sentence durations, durations of vowel (/i/) before word-initial stops, number of produced stop closures (/t/, /c/), durations of voicing in closures and VOT were measured. Preliminary analyses on three speakers show that they produced only one closure in most cases. Speakers occasionally produced two closures when slowing down their speaking rate. Vowels preceding /b, d, g/ were barely longer than vowels before /p, t, k/. When speakers produced one closure, closures were occupied by voicing for both /b, d, g/ and /p, t, k/. However, the average durations of voicing in closure were consistently longer for /b, d, g/ than for /p, t, k/. In addition, /b, d, g/ were produced in most cases with short-lag VOT, except when there were two closures, in which case, /b, d, g/ were produced with lead VOT. All /p, t,

k/ were produced with VOT greater than 30 ms. Articulatory studies are needed to assess how the sequence of stops is produced and whether the gestures for /t/ are completely or partially eliminated.

2pSC16. Modeling the effects of the lower airway on vowel spectra. Steven M. Lulich (MIT Speech Commun. Group, 77 Massachusetts Ave., MIT Bldg. 36-511, Cambridge, MA 02139, lulich@mit.edu)

Computational models of the acoustics of the lower airway (below the glottis) have been implemented by researchers in the speech and respiratory sciences. In general, these models are symmetrical and are studied independent of acoustic coupling to the vocal tract. We have constructed a new model that is asymmetrical and based on human morphometric data. The acoustic properties of the model are explored and compared with symmetrical models, and the effects of the model on vowel spectra are studied for varying amounts of lower airway and vocal-tract coupling via the glottis. Cochlear excitation patterns are also computed for the vowel spectra, and the effects of the lower airway acoustics on these excitation patterns are discussed.

2pSC17. Effect of tone on voice onset time in laryngeal and esophageal speech of Mandarin. Manwa Ng (Commun. Sci. and Disord., Long Island Univ., Brookville, NY 11548-1300) and Hanjun Liu (Northwestern Univ., Evanston, IL 60208)

Defined as the time between the release of burst of a stop and the onset of the following vowel, voice onset time (VOT) directly indicates the interarticulator timing between the voicing source and the articulatory movement in the vocal tract. A number of studies have investigated the VOT characteristics of different languages. Yet, few have studied how VOT is affected by tone level in a tone language, and by the type of phonation. The present study investigated the possible effect of tone level and phonation type on such interarticulator timing characteristics. VOT values were measured from the syllables /p^ha/, /t^ha/, and /k^ha/ produced at four Mandarin tone levels by eight laryngeal (NL) and seven esophageal (SE) speakers who were native speakers of Mandarin. Results indicated that significant differences in VOT were found in the NL speaker group, but not in the SE speaker group. With respect to phonation type, SE speakers showed significantly shorter VOT values than NL speakers. Such differences may be related to the use of pharyngoesophageal (PE) segment as a new sound source. SE speakers appear to take a shorter time to start PE segment vibration as compared to NL speakers.

2pSC18. Intrinsic vowel pitch in tracheo-esophageal speech. Michael Kiefte, Natalie Didrachnik, and Jillian Roswell (School of Human Commun. Disord., Dalhousie Univ., 5599 Fenwick St., Halifax, NS B3H 1R2 Canada)

The tendency for high vowels to have higher fundamental frequency (f_0) than low vowels, or intrinsic vowel pitch (IF_0), has been found in both laryngeal and esophageal speech [e.g., Gandour and Weinberg, Phonetica **25**, 140–164 (1980)]. In the present study, eight tracheo-esophageal speakers were asked to produce five repetitions of each of the four point vowels and f_0 was measured. Results showed significantly higher f_0 for high vowels than for low vowels, consistent with previous findings for laryngeal and esophageal speakers. In addition, f_0 was measured from two laryngectomees over a period of 4 months immediately following surgery to determine whether IF_0 is produced automatically or whether it develops over time. Neither of these subjects produced differences in IF_0 at any time within the 4-month period. These preliminary data appear to indicate that IF_0 is not an automatic consequence of either speech physiology or acoustics. More data from additional subjects will be presented. [Work supported by SSHRC.]

2pSC19. Computational aeroacoustics of flow through static model of human vocal tract. Jungsoo Suh, Steven H. Frankel, Luc Mongeau, and Michael W. Plesniak (School of Mech. Eng., Purdue Univ., West Lafayette, IN 47907)

Computational aeroacoustics studies of flow through the human vocal tract, here modeled as a planar channel with an orifice, hence referred to as the glottis, were conducted using large-eddy simulation (LES). Comparisons were made between LES predictions and experimental wall pressure measurements and particle-imaging-velocimetry flow fields. Rigid models of both converging and diverging glottal passages, each featuring a 20-deg included angle and a minimal glottal width of 0.04 cm with transglottal pressure of 15 cm H_2O, were studied. The compressible Navier-Stokes equations were accurately and efficiently integrated for the low Mach number flow through the use of an additive semi-implicit Runge-Kutta method and high-order compact finite-difference schemes for spatial discretization. Characteristic-based nonreflecting boundary conditions were used together with an exit zone in the context of a multiblock approach. Asymmetry of the flow due to a Coanda effect, and transition to turbulence were observed. An acoustic analogy based on the Ffowcs Williams-Hawkings equation was applied to decompose the near-field acoustic source into its monopole, dipole, and quadrupole contributions to assess glottal geometry effects on far-field sound. The results showed that dipole sources due to the unsteady forces exerted on the duct wall are dominant. [Work funded by NIH Grant RO1 DC03577.]

2pSC20. Modal analysis of a time-varying vocal tract with losses: A comparison of left and right eigenfunctions. Gordon J. Ramsay (Haskins Labs., 300 George St., New Haven, CT 06511-6624)

Sound propagation during speech can be modeled by a linear wave operator describing the spatiotemporal evolution of acoustic pressure and velocity distributions on a bounded domain representing the interior of the vocal tract. The eigenvalues of the wave operator determine the formant frequencies and bandwidths. The right eigenfunctions describe the standing wave patterns associated with each formant, whereas the left eigenfunctions describe how an arbitrary source distribution projects onto each formant. In speech research, it has always been implicitly assumed that left and right eigenfunctions are identical. This only holds if the wave operator is self-adjoint, which cannot be true when losses due to viscous damping, radiation effects, and wall vibration are taken into account. To examine this issue, a time-domain finite-volume simulation of acoustic wave propagation in the vocal tract was developed, including the above loss mechanisms. Eigenvalues and eigenfunctions were calculated from the resulting state-space recursion, for time-varying geometries representing VCV sequences containing fricatives and stops. Left and right eigenfunctions are similar in shape, but there are differences in the zero-crossing points and the amplitudes of the different lobes. Predictions based on right eigenfunctions alone will be incorrect. [Work supported by NIH.]

2pSC21. The interaction of linguistic and affective prosody in a tone language. Chinar Dara and Marc D. Pell (Commun. Sci. and Disord., McGill Univ., 1266 Pine Ave. West, Montreal, PQ H3G 1A8, Canada, chinar.dara@mail.mcgill.ca)

To address how a common set of acoustic properties of speech prosody modulate to convey linguistic and affective meanings concurrently, this study investigated the influence of phonemic tones on the expression of emotion (happy, sad, angry) and linguistic modality (declarative, interrogative) in a tone language, Punjabi. Base stimuli consisted of neutral sentences with either of the three contrastive tones in Punjabi (falling, level, high-rising) varying at the object position of the sentence only. Each of these sentences was elicited as a statement or a question for each of the three emotions by 6 native Punjabi speakers. Utterances were validated for adequate representation of the target emotion, and reliable exemplars for each of the emotions and linguistic modality were subjected to detailed acoustic analysis. Fundamental frequency and duration measures were analyzed for the stressed vowel of the keywords (subject, object, verb) and

the whole utterance for each of the tokens. Results demonstrated ways that the acoustic correlates of tone at the word level interact with that of the intonation and emotion at the sentence level, which were further compared with previous results on the effect of focus on intonation and emotional prosody [M. D. Pell, J. Acoust. Soc. Am. **109**, 1668–1680 (2001)].

2pSC22. The scope of effect of prosodic boundaries in articulation. Jelena Krivokapic (Dept. of Linguist., USC, 3601 Watt Way, GFS 301, Los Angeles, CA 90089-1693)

An articulatory study of the scope of prosodic boundary effects across boundaries of varying strengths is presented. Sentences in each of three prosodic conditions contain the following string: C1VC2VC3#VC4VC5VC6, where C is an alveolar consonant, V is a vowel, and # is a prosodic boundary. The boundaries are of three degrees of strength. Using articulator movement-tracking data (EMA) for the tongue tip articulator, consonant constriction formation and release duration, acceleration duration, and spatial magnitude are measured. Leftward and rightward temporal effects of the boundary on the consonants are investigated. Based on earlier studies (e.g. Cambier-Langeveld, Linguistics in the Netherlands, 13–24, 1997; Shattuck-Hufnagel and Turk, Proceedings 16th International congress on Acoustics and 135th Meeting Acoustical Society of America, 1235–1236, 1998; Berkovits, J. of Phonetics **21**, 479–489, 1993) and predictions of the prosodic gestural model (Byrd and Saltzman, J. of Phonetics, **31**, 149–180, 2003), it is expected that (1) the amount of articulatory lengthening will increase with boundary strength and (2) the degree of lengthening will decrease with distance from the boundary. Three subjects participated. The results show that at the boundary, gestures have longer constriction and acceleration durations, and the effect of the boundary increases with boundary strength, distinguishing up to three levels of boundary strength. The effect is local and diminishes with distance from the boundary. These results support the initial hypotheses. [Work supported by NIH.]

2pSC23. Acoustic and fiberscopic studies of the voice characteristics of school teachers under vocal abuse conditions with the loud voices of children. Noriko Kobayashi (School of Allied Health Sci., Kitsato Univ., 1-15-1 Kitasato, Sagamihara 228-8555 Japan, noriko@ahs.kitasato-u.ac.jp, Takashi Masaki, and Koichiro Nishiyama (Kitasato Univ, Sagamihara 228-8555, Japan)

It has been known that the incidence of vocal pathology was high in school teachers as they had to use a great amount of voice and loud phonation in noisy school settings with the children's loud voices. A proper vocal hygiene program and voice therapy are effective treatments for these cases. The final goal of the treatment should be to obtain the skills to use effective and healthy phonation even in noisy environments. In our study, three school teachers with vocal nodules, three speech therapists who were trained to produce nonconstricted voice even in noisy environments, and ten college students with no laryngeal pathology (control group) spoke eight sentences with one target word in two environments: a quiet environment and one with meaningful multi-talker babble (MMB). Acoustic analyses and fiberscopic examination revealed higher sound pressure and F0 levels and laryngeal constriction in MMB for the teachers before therapy and the normal talkers, but two speech therapists kept similar F0 levels and laryngeal conditions in both environments. After voice therapy, teachers produced less constricted voice under noise. These results suggested the efficacy of voice therapy with MMB for teachers to obtain effective and healthy voice in their noisy work environments.

2pSC24. Realizing question intonation in Mandarin with neutral tone. Fang Liu (Dept. of Linguist., Univ. of Chicago, 1010 E. 59th St., Chicago, IL 60637, liufang@uchicago.edu) and Yi Xu (Univ. College London, London, NW1 2HE, UK)

Previous research has found that postfocus pitch range is lowered in Mandarin even in questions, although the amount of lowering is smaller than in statements. This study investigates whether postfocus pitch range is raised in questions where the postfocus words carry the neutral tone, which, being conventionally considered targetless, may be better suited for manifesting intonation. Eight native speakers of Mandarin each produced 80 statements and 80 questions. The sentences ended either with a sequence of 5 neutral tones or with 3 neutral tones followed by 2 high tones. The sentences had two alternative focus locations, one on the full tone (high, rising, low, or falling) immediately preceding the neutral tone sequence, and the other on the low tone before the full tone. Results show that (1) postfocus lowering occurred in the neutral-tone-final sentences in both statements and questions, just as in the full-tone-final sentences, and (2) sentence-final neutral tone had falling F0 even in questions, thus contrasting with sentence-final high tone, which had rising F0 in questions. The findings are interpreted as indication that the neutral tone, being associated with weak articulatory effort, is not as effective as full tones in realizing the statement/question contrast.

2pSC25. Acoustic and articulatory correlates of contrastive focus. Melanie Thibeault, Lucie Menard, Annie Leclerc, Anthony Calabrino (Laboratoire de phonetique, UQAM, Case postale 8888, succursale Ctr. ville, Montreal QC, Canada, H3C 3P8, thibeault.melanie.2@courrier.uqam.ca), and Annie Brasseur (UQAM, Montreal QC, Canada, H3C 3P8)

This paper aims at examining production of French speakers in marking focus-induced prominence at the articulatory and acoustic levels. The corpus consisted of CVC syllables, where C corresponds to one of the stops /p t k/ and V is one of the vowels /i a u/. Target words were embedded in carrier sentences elicited in two prosodic conditions: neutral (unfocused) and under contrastive focus. Four adult speakers (all native speakers of French) pronounced ten repetitions of each sequence. The audio signal and tongue shapes were recorded using a digital camera and a SONOSITE 180 ultrasound. Formant frequencies, rms values, duration, and tongue contours corresponding to each vowel were extracted. Analyses show that prosodic context has a significant effect on acoustic and articulatory data for all vowels, increasing F1 and F2 under contrastive focus, compared to the neutral context. However, this stable acoustic pattern is achieved by various articulatory strategies across subjects. For example, 2 subjects produced /u/ under focus with a lower tongue body than in the neutral context, whereas the remaining 2 subjects had a higher tongue body in focused syllables. Results are compared to previous studies on articulatory and acoustic correlates of prosodic structure in French and English.

2pSC26. Rhythm in English clear speech. Rajka Smiljanic, Josh Viau, and Ann Bradlow (Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208, rajka@northwestern.edu)

This study investigates the effect of hyperarticulated, intelligibility-enhancing clear speech on English speech rhythm. Ramus *et al.* [Ramus *et al.*, "Correlates of linguistic rhythm in the speech signal," Cognition **72**, 265–292 (1999)] showed that temporal properties of a speech signal such as the percentage of vocalic intervals (%V) and variability of consonantal and vocalic intervals (C, V) can be related to phonological properties such as presence/absence of unstressed vowel reduction and syllable structure complexity, and are consequently quite successful at grouping languages into the traditional rhythmic classes (stress-, syllable-, and mora-timed). Here, we explore whether/how clear speech affects stress-timed characteristics of English sentences in terms of these measures. Results revealed that the proportion of vocalic intervals (%V) within sentences remained stable across speaking styles, i.e., consonants and vowels

were lengthened equally in clear speech. However, variability of both vocalic and consonantal intervals (V, C) was higher in clear than in conversational speech. The increase in V and C was achieved primarily through insertion/strengthening of short vocalic and consonantal segments that were dropped or coarticulated with surrounding sounds in conversational speech, making word and syllable boundaries more salient. These results suggest that increased intelligibility of clear speech can in part be attributed to prosodic structure enhancement by means of enhanced syllable and word boundaries demarcation.

2pSC27. Sensitivity of Mandarin speech to loudness-shifted voice feedback perturbation. Hanjun Liu, Qianru Zhang, Charles R. Larson (Northwestern Univ., 2240 Campus Dr., Evanston, IL 60208), and Yi Xu (Univ. College London, London, NW1 2HE, UK)

Recent studies have demonstrated that loudness-shifted voice auditory feedback leads to compensatory adjustments in voice amplitude during production of vowels. In this study we asked the question of whether Mandarin-speaking subjects would respond to loudness-shifted voice feedback during speech. Twenty-two native-speaking subjects with normal hearing and neurological background served as subjects. They said the Mandarin phrase ba ma, with the high tone on both syllables, as answers to three randomized prerecorded questions. The questions induced focus on the first or the second syllable, or on both syllables. During each production a ± 3 dB, 200- or 400-ms perturbation in voice feedback was presented 250 ms following voice onset. Randomized control trials were intermixed with test trials. Digitized voice records were converted to an rms trace representing voice amplitude. Traces were aligned with stimulus onset for event-related averaging. Results demonstrated that subjects responded to the loudness perturbation with compensatory adjustments in voice amplitude. The data extend recent findings indicating that voice feedback is used to help control amplitude during speech. [This work was supported by NIH Grant No. DC006243-01A1.]

2pSC28. Gradient alternations and gradient attraction. Katherine Crosswhite (Dept. of Linguist., Rice Univ., 6100 Main St., Houston, TX 77005, crosswhi@rice.edu)

Many phonological alternations have been shown to be sensitive to lexical ''gradient attraction,'' defined by Burzio [''Surface-to-surface morphology: When your representations turn into constrainsts,'' Rutgers Optimality Archive #341-0999 (http://roa.rutgers.edu) (1998)] as follows: (A) The overall structure of a word W (in both phonological and semantic components) is influenced by that of other words in the lexicon to which W is independently similar, and which are thought of as ''attractors'' of W. (B) Attraction is stronger where independent similarity is greater. Since most strong attractors will be morphologically related to W, this results in a limitation on phonological alternations that would produce allomorphy. So far, gradient attraction has been studied for categorical alternations, such as stress placement in English (cf. compArable versus cOmparable). However, many alternations previously believed to be categorical are in fact gradient [as reviewed in, e.g., Port and Leary, ''Against formal phonology,'' Language **81**, 927–964 (2005)]. This study documents the influence of lexical attraction on two gradient phenomena of English: polysyllabic shortening and flapping. Attractor strength was varied in two manners. First, each base used in the study was paired with both derivational and inflectional derivatives (i.e., rose-roses-rosy for polysyllabic shortening, and wait-waiting-waiter for flapping), and second by analyzing the effect of between-item variation in semantic relatedness for the derivationally related forms (cf. ice-icy versus nose-nosy).

2pSC29. Resonance tuning in soprano singing and vocal tract shaping: Comparison of sung and spoken vowels. Shrikanth Narayanan, Erik Bresch (USC Viterbi School of Eng., 3710 S. McClintock Ave, RTH 320, Los Angeles, CA 90089, shri@sipi.usc.edu), Stephen Tobin, Dani Byrd, Krishna Nayak, and Jon Nielsen (USC, Los Angeles, CA 90089)

Singing at high pitches by sopranos involves complex orchestration of vocal-fold mechanisms and vocal-tract shaping. A specific phenomenon is the tuning of vocal-tract resonances to the fundamental at high pitches. Measurement of vocal-tract resonances has shown that this is particularly true for the first resonance [Joliveau *et al.* (2004)]. Higher resonances show some increase in their values perhaps reflecting incidental effects rather than active tuning. We investigate vocal-tract shaping in soprano singing using a real-time MRI technique [Narayanan *et al.* (2004)] with synchronized noise-cancelled audio recording. A professionally trained Western-classical singer spoke and sang the vowels /a, i, u, o, e/ preceded by the consonant /l/ over an ascending scale (from B-flat below middle C, rising two octaves). Acoustic analysis shows clear evidence of F1 tuning to F0 in the higher-octave singing of the vowels. The vocal-tract shape was distinctly different at the higher pitches compared to vowels spoken at normal pitch and those sung at the lower octave, with fairly large oral and pharyngeal cavities, created by a highly bunched and retracted tongue, and a raised larynx. At the highest pitch, the vocal-tract shape differences between the vowels are considerably lessened, although some vestigial canonical tongue shapes appear to remain. [Work supported by NIH, USC Annenberg Center.]