Part V Language, Speech and Hearing

Section 1 Speech Communication

Chapter 1. Speech Communication

Academic and Research Staff

Professor Kenneth N. Stevens, Professor Jonathan Allen, Professor Morris Halle, Professor Samuel J. Keyser, Dr. Krishna K. Govindarajan, Dr. Helen M. Hanson, Dr. Joseph S. Perkell, Dr. Stefanie Shattuck-Hufnagel, Dr. Mario A. Svirsky, Dr. Alice Turk, Dr. Reiner Wilhelms-Tricarico, Dr. Yi Xu, Seth M. Hall

Visiting Scientists and Research Affiliates

Dr. Corine A. Bickley, Dr. Suzanne E. Boyce,[1] Dr. Anna Esposito,[2] Dr. Carol Espy-Wilson,[3] Dr. Robert E. Hillman,[4] Dr. Eva B. Holmberg,[4] Dr. Caroline Huang,[5] Dr. Harlan Lane,[6] Dr. John I. Makhoul,[7] Dr. Sharon Y. Manuel, Dr. Melanie L. Matthies,[3] Dr. Pascal H. Perrier,[8] Dr. David R. Williams,[9] Jane W. Wozniak[4]

Graduate Students

Hwa-Ping Chang, Marilyn Y. Chen, Karen Chenausky, Harold Cheyne, Jeung-Yoon Choi, Michael Harms, David M. Horowitz, Mark A. Johnson, Hong-Kwang J. Kuo, Sharlene A. Liu, Kelly L. Poort, Walter Sun, Lorin F. Wilde

Undergraduate Students

Max Chen, Howard Cheng, Laura C. Dilley, Genevieve Lada, Zöe Park, Edward D. Semper, Hemant Tanaja

Technical and Support Staff

Peter C. Guiod, Glenn S. LePrell, D. Keith North,[10] Arlene E. Wint

Sponsors

C.J. Lebel Fellowship
Dennis Klatt Memorial Fund
National Institutes of Health
- Grant F32-DC00194
- Grant F32-DC00205
- Grant P01-DC00361[11]
- Grant R01-DC00075
- Grant R01-DC00261[11]

National Institutes of Health
- Grant R01-DC00266[11]
- Grant R01-DC01291
- Grant R01-DC01925
- Grant R03-DC01721
- Grant R29 DC02525
- Grant T32-DC00038

---

1 Delrina, Boston, Massachusetts.

2 Department of Physics, Salerno University, Italy.

3 Boston University, Boston, Massachusetts.

4 Massachusetts Eye and Ear Infirmary, Boston, Massachusetts.

5 Altech Inc., Cambridge, Massachusetts, and Boston University, Boston, Massachusetts.

6 Department of Psychology, Northeastern University, Boston, Massachusetts.

7 Bolt, Beranek and Newman, Inc., Cambridge, Massachusetts.

8 Institut de Communication Parlé Grenoble, France.

9 Sensimetrics Corporation, Cambridge, Massachusetts.

10 Deceased February 9, 1995.

11 Under subcontract to Massachusetts Eye and Ear Infirmary.

National Science Foundation
Grant INT 94-21146[12]
Grant IRI 89-05439[13]

1.1 Studies of Normal Speech Production

1.1.1 Speech Motor Control: Phonemic Goals and the Use of Feedback

Two oral presentations contained the following theoretical overview of speech motor control along with examples of supporting data from studies of articulatory-to-acoustic motor equivalence and changes in the hearing status of cochlear implant and NF2 patients.

Much of the information content in speech is conveyed by phonemic mechanisms, which consist of neuro-muscular synergisms that achieve segmental articulatory and acoustic goals. Some of the goals are determined by quantal (nonlinear) relations between articulation and sound. The goals may also be influenced by other principles, such as a compromise between sufficient perceptual contrast and economy of articulatory effort. This leads to the prediction that the goal definitions correspond to regions (as opposed to points) in acoustic and articulatory space. Thus the goals are characterized by some parameter variation, which is possible in part because listeners can understand variable speech. The speech motor control system includes a number of phonemic mechanisms, organized in a modular, hierarchical fashion.

When utterances are produced, sequences of goal specifications are converted to smooth, appropriately-timed articulatory movements by the control system, which takes into account the biomechanical properties of the articulators. To help keep acoustic variability within perceptually-acceptable limits, speech motor control mechanisms may include a strategy by which different parts of the vocal-tract area function are adjusted in a complementary ("motor equivalent") manner. This strategy takes advantage of the fact that for some sounds, a similar acoustic transfer function can be achieved with somewhat different area functions. The existence of such a strategy and the idea that speech motor programming is based in part on acoustic goals are supported by data that show motor equivalent relations between lip and tongue constrictions in production of the sounds /u/, /r/ and /š/. The strategy also implies that speech motor programming relies on the use of a robust internal model for feedforward control. The internal model represents relations between speech motor commands and the sound output.

Information about the relation between speech motor commands and the sound output is established during speech acquisition with the use of auditory feedback. The role of auditory feedback for speech production in adults is unclear, since the speech of people deafened as adults can remain intelligible for decades. However, their speech often develops abnormalities, indicating some role for auditory input in adult speech motor control. Studies of changes in speech production when postlingually-deafened patients receive cochlear implants have led us to hypothesize that auditory feedback has at least two functions in adult speech motor control. These are: (1) maintenance of the information represented in the internal model, and (2) monitoring the transmission channel to help make situation-dependent adjustments in "postural" settings of parameters that underlie average sound level, speaking rate, voice fundamental frequency, and vowel formants, which influence clarity and intelligibility. By inference, phonemic settings should be less labile than postural settings. Since phonemic settings and postural settings affect the same articulators, there can be interactions between them, but in some cases their changes can be observed separately.

1.1.2 Lip Protrusion Movements in Normal versus Clear Speech

A kinematic analysis of midsagittal movement (EMMA) data was performed for upper lip protrusion from /i/ to /u/ in /iku/, /ikku/ and /iktku/, produced by a speaker of American English. Utterances were embedded in a carrier sentence and spoken in four conditions: normal rate, normal rate and clear, fast rate, fast rate and clear. The results show several regularities across all conditions: (1) the delay between the protrusion onset and the acoustic onset of /u/ varies linearly with the vowel-to-vowel duration (VVD), /i/-end to /u/-onset, with a slope of 0.7; (2) the overlap between the protrusion movement and vowel /i/ decreases as VVD increases, with no significant differences between

[12] U.S.-India Cooperative Science Program.

[13] Under subcontract to Boston University.

conditions for similar VVD values; (3) the relationship between the ratio of maximum velocity (Vmax) to protrusion amplitude (Amp) and the movement duration (T) is constant: Vmax/Amp=c/T; (4) the number of velocity peaks increases linearly with VVD. These regularities suggest that the same control strategy underlies the protrusion gesture, whatever the condition. However, for a given rate, clear tokens were produced with longer VVDs and larger protrusion amplitudes. For similar VVD values, fast clear tokens had larger values of Vmax than normal tokens.

1.1.3 Physiological Modeling of Speech Production

For a computational finite element tongue model, a new finite element type was implemented experimentally, using Matlab and C subroutines. The code computes triquadratic interpolation of the displacements from 27 nodes per element, and a pressure field that is represented as linear function with four coefficients. All computations are vectorized. The total system of equations constitutes a mixed differential and algebraic system. The differential equations result from the dynamics of the system, and the purpose of the algebraic equations is to determine, for each element, four coefficients of a pressure field, such that each element's incompressibility condition is maintained. To solve the large system of equations, the previously used explicit method is being replaced by an implicit algorithm that is based on the conjugate gradient method.

One goal of the biomechanical vocal tract modeling is to emulate the morphological characteristics of individual speakers. For this purpose, MRI data are used to obtain data on the individual speaker, beginning with the extraction of vocal tract shapes (area functions). Software has been developed that loads three sets of MRI data, representing sagittal, coronal, and transversal series of slices of the oral and pharyngeal region of a subject who maintains an articulatory posture for a vowel or consonantal constriction. The spatial locations and spatial ranges of the three data sets are adjusted so that, at any spatial location, maximal similarity is achieved. The adjustment method is semi-automatic: initial values are set interactively using visual feedback; then an iterative search tries to maximize a quantitative measure of similarity (a binary correlation measure, formed as the relative number of equal bits in two thresholded images). The next stage extracts connected regions in the three-dimensional stacks of MRI images and eliminates the smallest (spurious) regions. For the final step, software has been implemented for regions growing in three dimensions and for hand-editing binary images to aid the extraction of the vocal tract shapes. MRI data have been obtained on several speakers. Data from MRIs helps us to make substantial progress in obtaining high quality images.

1.1.4 Improvements in Facilities for Speech Production Research

A number of improvements were made in hardware and software facilities for gathering and analyzing data on articulatory movements. We received a new set of electronics for our EMMA system and new, smaller, pre-wired transducers. The new components were tested extensively, and a number of steps were taken to assure adequate performance. A new procedure was developed to facilitate the mounting of transducers on subjects' articulators. An interface box was designed and constructed to increase the channel capacity of the EMMA real-time display and eliminate interference between the real-time display and output voltages of the EMMA system. Improvements were also made in the functionality of the program for the EMMA real-time display. Several improvements were made in methods for taking and analyzing video recordings of subjects' lip areas. Additional algorithms were developed and used for the interactive, semi-automatic extraction of kinematic and timing data from movement signals.

1.2 Speech Research Relating to Special Populations

1.2.1 Speech Production of Cochlear Implant Patients

Intersyllabic Regulation of F0 and SPL Both With and Without Auditory Feedback

Syllable-to-syllable fluctuation of *F0* and SPL were measured in readings of a passage by (1) four post-lingually deafened adults, recorded before and after they received cochlear implants, and (2) one adult with neurofibromatosis-2 (NF2), who was initially profoundly deaf in one ear and had a mild to moderate hearing loss in the other (aided). Three of the four cochlear implant users reliably reduced syllable-to-syllable fluctuation in *F0* and SPL following the activation of their speech processors. The fourth speaker began with and maintained the regularity achieved by the others post-activation. In recordings made following the onset of bilateral profound deafness, the NF2 subject showed increased syllable-to-syllable fluctuation in *F0* and SPL.

Results from another experiment[14] in which multiple repetitions of vowels in an /hVd/ context were produced by cochlear implant users with their processors turned off and on, suggest that some subjects showed less token-to-token variability in *F0* and SPL with their processors turned on. The present results indicate that auditory feedback may also constrain the syllable-to-syllable variability of *F0* and SPL contours.

Acoustic and Articulatory Measures of Sibilant Production Both With and Without Auditory Feedback from a Cochlear Implant

The articulator positions of a subject with a cochlear implant were measured with an electro-midsagittal articulometer (EMMA) system both with and without auditory feedback available to the subject via his implant. Acoustic analysis of sibilant productions included specific measures of their spectral properties as well as the *F3* formant amplitude. More general postural characteristics of the utterances such as speech rate and sound level were also measured. Because of the mechanical and aerodynamic interdependence of the articulators, the postural variables must be considered before attributing speech improvement to the selective correction of a phonemic target with the use of auditory feedback. The tongue blade position was related to the shape and central tendency of the /š/. spectra; however, changes in the spectral contrast between /s/ and /š/. were not related to changes in the more general postural variables of rate and sound level. These findings suggest that the cochlear implant is providing this subject with important auditory cues that he can use to monitor his speech and maintain the phonemic contrast between /s/ and /š/.

1.2.2 Acoustic and Perceptual Effects of Endoscopic Nasal Surgery

Past research has shown that some modifications occur in the speech of patients who have undergone endoscopic nasal surgery. These changes have been measured in the spectra of nasal consonants and nasal vowels produced by these individuals. The measurements suggest that the surgery has two effects on the nasal vowels: an increased prominence of spectral peaks, due presumably to the decreased acoustic losses in the modified nasal cavity, and an enhanced influence of sinus resonances in the spectrum. Two of the nasal vowels produced by the patients were excised from the speech and were presented to listeners, who compared the nasality in the vowels recorded pre- and post-operatively. The vowel /i/ was judged to be less nasal after the surgery and /æ/ sounded more nasal. A theoretical account of this result has been proposed in terms of the two abovementioned effects.

1.2.3 Speech Input for Dysarthric Computer Users

Previous work on this project has reported on the kinds of speech errors made by a group of dysarthric speakers. Errors or inconsistencies in production were especially prevalent for obstruent consonants. Within this class of sounds, the greatest difficulties were experienced for fricatives and affricates produced by forming a constriction with the tongue blade. One goal of this project is to facilitate interaction with computers by these individuals through the use of a speech recognizer. Based on a study of the characteristics of the speech sounds produced by each speaker, a list of words was selected for the speaker to control the computer. The words were designed to avoid the sounds that the speaker was unable to produce consistently. The percent error in recognition of these words for different speakers was 10-35 percent less than the error when the words were randomly selected.

1.2.4 Voice Source Analysis of Speakers with Vocal-Fold Nodules

The purpose of this research is to study the voice production mechanism in patients with vocal-fold nodules in an attempt to understand how they happen. Vocal-fold nodule formation has often been thought to be caused by voice misuse or abuse. This condition can cause partial or total loss of voice and often requires treatment with voice therapy and/or surgery.

Aerodynamic measures for a selected set of five female nodule patients and control subjects were studied and related to published data. The subjects were asked to speak at three levels of voice effort: comfortable, intermediate, and loud. In general, the nodule patients were found to be speaking louder when speaking at comfortable voice effort than the normals. They used increased transglottal pressure, and the glottal flow had a greater AC compo-

[14] M.A. Svirsky, H.L. Lane, J.S. Perkell, and J.F. Wozniak, *J. Acoust. Soc. Am.* 92(3): 1284-1300 (1992).

nent, maximum flow declination rate, and minimum flow.

One nodule patient exhibiting high transglottal pressure and maximum flow declination rate was chosen for more detailed study. This patient had undergone voice therapy. The aerodynamic and acoustic properties of this patient's speech before and after voice therapy were compared, and in general showed a change toward more normal characteristics. Voice therapy appeared to reduce the transglottal pressure used by this patient and increase the speech intensity for a given pressure. The speech before therapy sometimes underwent large changes in amplitude during sustained vowel phonation, and spectral measures show that some of the low amplitude regions were more breathy in character or accompanied by diplophonia in severe cases when the voice "catches." All observations made have been on a small scale and need to be verified for other recordings for this subject as well as other patients who have the same voice pathology.

1.3 Studies of the Acoustics, Perception, and Modeling of Speech Sounds

1.3.1 Effects of Tracheal Resonances on Vowel Spectra

We often observe prominences in the spectra of vowels that are not attributable to natural frequencies of the vocal tract. These irregularities in the spectrum result from acoustic coupling, through the glottis, between the vocal tract and the trachea, and can be modeled as pole/zero pairs introduced into the transfer function between the glottal source and the output at the mouth. The amount of separation between a pole and a zero in such a pair depends on the size of the glottal opening, i.e., the glottal impedance. For different speakers, the average glottal area, and hence the magnitude of the glottal impedance may be different. Thus, tracheal resonances should be more evident in vowel spectra for individuals who phonate with a glottis that remains partially open throughout a glottal cycle. Such individuals are known to exhibit a greater high-frequency tilt in

the glottal spectrum and a greater first-formant bandwidth due to increased acoustic losses at the partially open glottis.

Theoretical analysis suggests that there are two kinds of acoustic evidence for acoustic coupling to the trachea: one is the presence of spectral prominences in addition to the prominences due to vocal-tract resonances, or formants, and the other is the disruption of prominences due to formants as they pass through frequency ranges of tracheal resonances. Acoustic data obtained from vowels produced by 22 female speakers were studied for such evidence of acoustic coupling to the trachea, and were examined in relation to acoustic measures of average glottal opening. The results show that when spectral measures indicate significant glottal openings, evidence for tracheal resonances appear in vowel spectra. The effect of the tracheal resonances on the spectrum is more pronounced for speakers who phonate with a greater average glottal opening.

1.3.2 Estimating the Effects of Acoustic Losses in the Nasal Cavity

Creation of a velopharyngeal opening during vowel production has two major acoustic consequences. One consequence is to shift the frequencies of the vocal-tract resonances or formants and to introduce additional resonances due to the nasal cavities or sinuses. Another influence is to create additional acoustic losses because of the relatively large area of the surfaces of the nasal passages. A theoretical examination of the influence of these losses on the bandwidth of the first formant has been carried out as part of an effort to develop acoustic measures of nasalization. The theory uses as a starting point published acoustic data on the transfer function of the nasal cavity. From these data, the impedance looking into the nasal cavity can be estimated, particularly the resistive component that accounts for the losses. This impedance is connected across the main vocal tract at about its midpoint. The effect of the resistive component of this impedance on the bandwidth of the first formant can then be calculated. The calculations show that the losses in the nasal cavity can lead to a 10 dB decrease in the amplitude of the first-formant peak when there is an appreciable velopharyngeal opening. This decrease in *F1* amplitude is consistent with acoustic measurements comparing nasal and nonnasal vowels, although some variability in this measure is observed. Perceptual studies with synthetic speech have shown that this increased *F1* bandwidth is a major contributor to the perception of nasality.

1.3.3 Lateral Consonants

As part of a continuing study of liquid consonants, we have been comparing the acoustic correlates of syllabic / l̩ / (in words like **buckle** and the unstressed syllable nucleus /o/ (in words like **bucko**). We have completed the analysis of several such utterances by four speakers. The principal findings from the analysis are a decreased prominence for the second formant and an increased prominence for the third formant for / l̩ / although there are inividual differences in the magnitude of these effects. Perceptual experiments with synthetic speech show that both of these spectral changes are cues for syllabic / l̩ / with the second-formant bandwidth being the stronger cue. The increased *F2* bandwidth can be explained through the analysis of the behavior of models of / l̩ / production, which show that acoustic losses arise from airflow through a narrow constriction formed by the tongue blade.

1.4 Models of Lexical Representation and Lexical Access

1.4.1 Theoretical Background: Representation of Segments in the Lexicon

In a proposed model for lexical access that is under development, lexical items are stored in memory in terms of sequences of segments, each of which is specified as a set of features that are arranged in a hierarchical manner. We have proposed an arrangement of features in a tree-like structure that has three principal nodes which dominate nested sets of articulators. The root node at the top of the tree dominates the production of vowels. The next node in the hierarchy dominates the production of glides, and the lowest node (the supranasal node) dominates nodes designating articulators involved in the production of consonant sounds. The major class features [consonantal] and [vocalic] are specified by designating which of the three primary nodes is dominant. The remaining articulator-free features [continuant] and [strident] are distinctive only for consonants, and they are attached to the node designating the primary articulator for the consonant.

1.4.2 Identification of Landmarks

One component of a proposed model of lexical access is a procedure for locating a sequence of landmarks that identify times at which consonantal closures and releases are produced. The process of locating these landmarks also leads to specification of manner features such as [continuant] and [sonorant] for the consonants.

A consonantal landmark detector has been developed, and its performance has been evaluated for consonants in a variety of contexts, for speech in noise, and for telephone-quality speech. The detector showed relatively few errors for wideband noise-free speech but was more error-prone for consonants occurring in a context with a preceding stressed vowel and a following reduced vowel—a context in which the landmarks are less salient. The detection error rate rose significantly for a signal-to-noise ratio (with speech-shaped noise) of 10 dB or less and also more than doubled with telephone speech relative to wide-band noise-free speech. Improved performance was obtained if the landmark-finding algorithm was adjusted for the specific noise or band-limited condition.

The development of methods for locating consonantal landmarks is being extended to include the detection of landmarks for glide consonants. These consonants do not create abrupt discontinuities in the spectrum, and the landmarks are defined as points in the sound where the vocal tract is maximally constricted for the glide. Algorithms have been developed for locating local minima in amplitude and in first-formant frequency. These algorithms are being evaluated with several databases of utterances, including a group of sentences that have been designed specifically to evaluate landmark detection for glides in various phonetic contexts.

1.4.3 Automatic Measurement of Acoustic Correlates of Place of Articulation for Stop Consonants

In order to identify the distinctive features of a consonant, the acoustic signal must be characterized by an informative set of measurements, which can then be classified. Based on suggestions in the phonetics literature as well as theoretical analysis of consonant production, we have designed a set of acoustic measurements to extract information about the place of articulation of a stop or nasal consonant release. Algorithms were developed to collect four burst shape measurements at a stop release, two measurements of spectral change at a nasal release, and temporal samples of the first three

formant trajectories. These algorithms were designed and optimized manually, and then re-adjusted using a gradient descent algorithm to give minimum classification error on a large data set. Measurements were evaluated in the acoustic domain, and burst spectral measurements were also transformed to a related articulatory domain. Consonant and vowel-dependent shifts predicted by theory were observed in the mean and variance of the inferred articulatory variables.

1.5 Speech Production Planning and Prosody

Our work in speech production planning this year focused on three aspects of our goal of developing an integrated model of segmental and prosodic aspects of planning: (1) development of a corpus of tape-recorded speech errors with special focus on sublexical interaction errors between two sites in an utterance (e.g., phonemic spoonerisms such as "cry to teep" for "try to keep", and anticipatory/perseveratory substitutions such as "pound the focket" for "found the pocket"), and prosodic annotation of this corpus; (2) development and analysis of a corpus of professionally read, nonprofessionally read and spontaneous speech, with prosodic annotation, to permit analysis of glottalization and other phonetic markers of prosodic structure, and (3) continuing expansion and analysis of a prosodically labeled database of professional FM radio newscast speech. Results from analyses of these databases include:

1. Glottalization of vowel-initial syllables, previously shown to occur at prosodically significant locations in read speech, shows a similar pattern in spontaneous dialogues, occurring at the onset of intermediate intonational phrases as well as in conjunction with pitch accents. Although individual speakers differ substantially in their overall glottalization rate, each shows the same tendency to glottalize vowel-initial syllables more often at these prosodically significant locations.

2. Moreover, for most speakers a set of additional acoustic features often accompany this vowel-initial glottalization, including breathy onset, low amplitude and low *F0*. Re-analysis of tokens labeled as nonglottalized showed that these features can occur on their own at the onset of vowel-initial words. When they occur, it is at almost entirely at the same prosodically significant locations, i.e., at phrase onsets and prominent syllables. In conjunction with other work that is being reported from other laboratories, these findings support the hypothesis that the onset of a prosodic constituent is strengthened.

3. Following up on earlier analyses of early pitch accent within late-stress words like "financial" and "Massachusetts", we found that early pitch accent occurs more consistently for alternating stress words like "Massachusetts" than for adjacent stress words like "financial". We are currently investigating the implications of this result for the role of branching foot structure in pitch accent assignment in English.

1.5.1 Progress in other Aspects of Speech Production Planning

1. We are transferring our corpus of 15,000 spoken errors from index cards to a computerized database, using a categorization system developed cooperatively with Merrill Garrett and Victoria Fromkin. The database will enable rapid searches for errors of certain types, facilitate the testing of large numbers of hypotheses to make it possible to share subsets of errors with the research community in presorted form, reducing the chances that error data will be miscategorized and misinterpreted.

2. We participated in the Workshop on Prosodic Transcription which preceded the International Congress of Phonetic Sciences in Stockholm in August 1995. We also began a collaborative effort which we hope will lead to an extension of our analysis of glottalization of vowel-initial syllables as a function of prosodic structure to examples in a prosodically labeled database of spoken German.

1.6 Development of Computer and Data Processing Facilities

We have completed the port of the Klatt Analysis Tools to the X11 environment and the integration of the Klatt speech synthesis utility, KLSYN, into the overall suite of Klatt Tools. This work completes a total integration of all the Klatt Tools used by the Speech Group into one seamless application package (called XKL), which is now available to the general research community in the X11 distributed graphics environment. Further ports of XKL have been done for the Sun and Silicon Graphics workstation environments, and a Project Athena port is in progress. The RLE Speech Group has continued to make increasing use of the new generation of very high performance RISC architecture UNIX workstations for signal analysis and biomechanical

modeling, as well as general purpose data analysis using MATLAB, PERL and the Entropics tool package. Work has begun on the migration of the speech physiology VAX-based real-time data acquisition and stimulus presentation hardware and software to a Pentium-based PC platform.

1.7 Publications

1.7.1 Published Papers

Chen, M. "Acoustic Parameters of Nasalized Vowels in Hearing-impaired and Normal-hearing Speakers." *J. Acoust. Soc. Am.* 98 (5): 2443-2453 (1995).

Dilley, L., and S. Shattuck-Hufnagel. "Variability in Glottalization of Word Onset Vowels in American English." *Proceedings of the International Congress of Phonetic Sciences,* Stockholm, Sweden, 1995, vol. 4, pp. 586-589.

Hanson, H.M. "Individual Variations in Glottal Characteristics of Female Speakers." *Proceedings of the IEEE ICASSP,* Detroit, Michigan, 1995, vol. 1, pp. 772-775.

Hanson, H.M., and K.N. Stevens. "Sub-glottal Resonances in Female Speakers and their Effect on Vowel Spectra." *Proceedings of the International Congress of Phonetic Sciences,* Stockholm, Sweden, 1995, vol. 3, pp. 182-185.

Holmberg, E.B., R.E. Hillman, J.S. Perkell, and S. Goldman. "Comparisons Among Aerodynamic, Electroglottographic, and Acoustic Spectrum Measures of Female Voices." *J. Speech Hear. Res.* 38: 1212-1223 (1995).

Lane, H., J. Wozniak, M.L. Matthies, M.A. Svirsky, and J.S. Perkell. "Phonemic Resetting vs. Postural Adjustments in the Speech of Cochlear Implant Users: An Exploration of Voice Onset Time." *J. Acoust. Soc. Am.* 98: 3096-3106 (1995).

Liu, S.A. "The Effect of Vowel Reduction on Landmark Detection." *Proceedings of the International Congress of Phonetic Sciences,* Stockholm, Sweden, 1995, vol. 4, pp. 136-139.

Manuel, S.Y. "Speakers Nasalize /ð/ after /n/, But Listeners Still Hear /ð/." *J. Phonetics* 23: 453-476 (1995).

Manuel, S.Y., and K.N. Stevens. "Formant Transitions: Teasing Apart Consonant and Vowel Contributions." *Proceedings of the International Congress of Phonetic Sciences,* Stockholm, Sweden, 1995, vol. 4, pp. 436-439.

Perkell, J.S., J. Manzella, J. Wozniak, M. Matthies, H. Lane, M. Svirsky, P. Guiod, L. Delhorne, P. Short, M. MacCollin, and C. Mitchell. "Changes in Speech Production Following Hearing Loss Due to Bilateral Acoustic Neuromas." *Proceedings of the International Congress of Phonetic Sciences,* Stockholm, Sweden, 1995, vol. 3, pp. 194-197.

Perkell, J.S., M.L. Matthies, M.A. Svirsky, and M.I. Jordan. "Goal-based Speech Motor Control: A Theoretical Framework and Some Preliminary Data." *J. Phonetics* 23: 23-35 (1995). Also in *Disorders of Motor Speech: Assessment, Treatment, and Clinical Characterization.* Baltimore: Paul H. Brookes, 1995, pp. 27-42.

Poort, K.L. "Stop Consonant Production: An Articulation and Acoustic Study." *Proceedings of the International Congress of Phonetic Sciences,* Stockholm, Sweden, 1995, vol. 3, pp. 444-447.

Shattuck-Hufnagel, S. "Pitch Accent Patterns in Adjacent-stress vs. Alternating-stress Words in American English." *Proceedings of the International Congress of Phonetic Sciences,* Stockholm, Sweden, 1995, vol. 3, pp. 656-659.

Shattuck-Hufnagel, S. "The Importance of Phonological Transcription in Empirical Approaches to Stress Shift vs. Early Accent." In *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence.* Eds. B. Connell and A. Arvaniti. Cambridge: Cambridge University Press, 1995, pp. 128-140.

Sluijter, A.M.C., S. Shattuck-Hufnagel, K.N. Stevens, and V.J. van Heuven. "Supralaryngeal Resonance and Glottal Pulse Shape as Correlates of Stress and Accent in English." *Proceedings of the International Congress of Phonetic Sciences,* Stockholm, 1995, vol. 2, pp. 630-633.

Stevens, K.N. "Models of Speech Production and Perception." In *European Studies in Phonetics and Speech Communication.* Eds. G. Bloothooft, V. Hazan, D. Huber, and J. Llisterri. The Hague: CIP-Gegeens Koninklijke Bibliotheek, 1995, pp. 52-57.

Stevens, K.N. "Applying Phonetic Knowledge to Lexical Access." *Proceedings of the Fourth*

European Conference on Speech Communication and Technology, Madrid, Spain, 1995, vol. 1, pp. 3-1.

Wilde, L. "Quantifying Time-varying Spectra of English Fricatives." *Proceedings of the International Congress of Phonetic Sciences,* Stockholm, Sweden, 1995, vol. 4, pp. 120-123.

Wilhelms-Tricarico, R. "Physiological Modeling of Speech Production: Methods for Modeling Soft-tissue Articulators." *J. Acoust. Soc. Am.* 97(5): 3085-3098 (1995).

Wilhelms-Tricarico, R. and J.S. Perkell. "Biomechanical and Physiologically-based Speech Modeling." *Proceedings of the International Congress of Phonetic Sciences,* Stockholm, 1995, vol. 2, pp. 68-75.

Xu, Y. "The Effect of Emphatic Accent on Contextual Tonal Variation." *Proceedings of the International Congress of Phonetic Sciences,* Stockholm, Sweden, 1995, vol. 3, pp. 668-671.

1.7.2 Papers Accepted for Publication

Halle, M., and K.N. Stevens. "The Postalveolar Fricatives of Polish." *Festschrift for Osamu Fujimura,* Mouton de Gruyter. Forthcoming.

Hillman, R.E., E.B. Holmberg, J.S. Perkell, J. Kobler, P. Guiod, C. Gress, and E.E. Sperry. "Speech Respiration in Adult Females with Vocal Nodules." *J. Speech Hear. Res.* Forthcoming.

Lane, H., J. Wozniak, M.L. Matthies, M.A. Svirksy, J.S. Perkell, M. O'Connell, and J. Manzella. "Regularity of Fundamental Frequency and Amplitude Contours with and Without Auditory Feedback." *J. Acoust. Soc. Am.* Forthcoming.

Matthies, M.L., M.A. Svirsky, J.S. Perkell, and H. Lane. "Acoustic and Articulatory Measures of Sibilant Production with and without Auditory Feedback from a Cochlear Implant." *J. Speech Hear. Res.* Forthcoming.

Perkell, J.S. "Properties of the Tongue Help to Define Vowel Categories: Hypotheses Based on Physiologically-oriented Modeling." *J. Phonetics.* Forthcoming.

Perkell, J.S. "Articulatory Processes." In *Handbook of Phonetic Sciences.* Eds. W. Hardcastle and J. Laver. Oxford, England: Blackwell. Forthcoming.

Perkell, J.S., and M.H. Cohen. "Token-to-token Variation of Tongue-body Vowel Targets: The Effect of Context." *Festschrift for Osamu Fujimura,* Mouton de Gruyter. Forthcoming.

Shattuck-Hufnagel, S., and A. Turk. "A Tutorial on Prosody for Investigators of Auditory Sentence Processing." *J. Psycholinguistic Res.* Forthcoming.

Stevens, K.N. "Models of Speech Production." In *Handbook of Acoustics,* Ed. M. Crocker. New York: Wiley. Forthcoming.

Stevens, K.N. "Articulatory-acoustic-auditory Relationships." In *Handbook of Phonetic Sciences.* Eds. W. Hardcastle and J. Laver. Oxford, England: Blackwell. Forthcoming.

Stevens, K.N. "Critique: Articulatory-acoustic Relations and their Role in Speech Perception." *J. Acoust. Soc. Am.* Forthcoming.

Wilhelms-Tricarico, R. "Biomechanical and Physiologically-based Speech Modeling." *J. Phonetics.* Forthcoming.

1.7.3 Papers Submitted for Publication

Chen, M.Y. "Baseline Acoustic Correlates of English and French Nasalized Vowels." Submitted to *J. Acoust. Soc. Am.*

Dilley, L., S. Shattuck-Hufnagel and M. Ostendorf. "Glottalization of Vowel-Initial Syllables as a Function of Prosodic Structure." Submitted.

Hanson, H.M. "Glottal Characteristics of Female Speakers: Acoustic Correlates." Submitted to *J. Acoust. Soc. Am.*

Shattuck-Hufnagel, S. "Acoustic-phonetic Evidence for Prosodic Structure: A Review." An invited contribution to a festschrift.

Svirsky, M.A., K.N. Stevens, M.L. Matthies, J. Manzella, J.S. Perkell, and R. Wilhelms-Tricarico. "Tongue Surface Displacement during Obstruent Stop Consonants." Submitted to *J. Acoust. Soc. Am.*

Theses

Chang, H.-P. *Speech Input for Dysarthric Computer Users.* Ph.D. diss. Dept. of Mechan. Eng., MIT, 1995.

Chen, M.Y. *Acoustic Correlates of Nasality in Speech.* Ph.D. diss. Harvard-MIT Division of Health Sciences and Technology, 1995.

Hanson, H.M. *Glottal Characteristics of Female Speakers.* Ph.D. diss. Division of Applied Sciences, Harvard University, 1995.

Liu, S.A. *Landmark Detection for Distinctive Feature-Based Speech Recognition.* Ph.D. diss. Dept. of Electr. Eng. and Comput. Sci., MIT, 1995.

Poort, K.L. *Stop Consonant Production: An Articulation and Acoustic Study.* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1995.

Wilde, L.F. *Analysis and Synthesis of Fricative Consonants.* Ph.D. diss. Dept. of Electr. Eng. and Comput. Sci., MIT, 1995.