Speech Communication

Sponsors
C.J. Lebel Fellowship
Dennis Klatt Memorial Fund
Donald North Memorial Fund
National Institutes of Health (Grants R01-DC00075,
R01-DC01291,
R01-DC01925,
R01-DC02125,
R01-DC02978,
R01-DC03007,
1 R29 DC02525,
T32-DC00038,
and National Science Foundation (Grants INT-9615380 (US-France Cooperative Research), and INT-9821048 (US-Germany Cooperative Research))

Academic and Research Staff
Professor Kenneth N. Stevens, Professor Jonathan Allen, Professor Morris Halle, Professor Samuel J. Keyser, Dr. Joseph S. Perkell, Dr. Stefanie Shattuck-Hufnagel, Dr. Marilyn Chen, Dr. Jeung-Yoon Choi, Dr. Mark Tiede, Dr. Reiner Wilhelms-Tricarico, Dr. Lisa Lavoie, Jennell Vick, Majid Zandipour, Seth Hall, John Gould.

Visiting Scientists and Research Affiliates
Dr. Takayuki Arai, Department of Electrical and Electronics Engineering, Sophia University, Tokyo, Japan.
Dr. Corine A. Bickley, Department of Communication Sciences and Disorders (IHP), Massachusetts General Hospital, Boston, Massachusetts, and Voice Services Division, Comverse, Cambridge, Massachusetts.
Dr. Suzanne E. Boyce, Department of Communication Disorders, University of Cincinnati, Cincinnati, Ohio.
Dr. Carol Y. Espy-Wilson, Department of Electrical Engineering, Boston University, Boston, Massachusetts.
Dr. Krishna Govindarajan, SpeechWorks International, Boston, Massachusettts.
Dr. David Gow, Department of Psychology, Salem State College, Salem, Massachusetts, and Department of Neuropsychology, Massachusetts General Hospital, Boston, Massachusetts.
Dr. Frank Guenther, Department of Cognitive and Neural Systems, Boston University, Boston, Massachusetts.
Dr. Helen M. Hanson, Sensimetrics Corporation, Somerville, Massachusetts.
Dr. Robert E. Hillman, Mass Eye and Ear Infirmary, Boston, Massachusetts.
Dr. Caroline Huang, SpeechWorks International, Boston, Massachusettts.
Aaron Im, Department of Neuropsychology, Massachusetts General Hospital, Boston, Massachusetts.
Dr. Harlan Lane, Department of Psychology, Northeastern University, Boston, Massachusetts.
Dr. Sharon Y. Manuel, Department of Speech Language Pathology & Audiology, Northeastern University, Boston, Massachusetts.
Dr. Melanie Matthies, Department of Communication Disorders, Boston University, Boston, Massachusetts.
Dr. Richard McGowan, Sensimetrics Corporation, Somerville, Massachusetts.
Dr. Rupal Patel, Department of Bio-behavioral Studies, Teachers College Columbia University, New York.
Dr. Alice Turk, Department of Linguistics, University of Edinburgh, Edinburgh,
United Kingdom.
Dr. Nanette Veilleux, Department of Computer Science, Simmons College, Boston, Massachusetts.
Dr. Lorin Wilde, Lernout & Hauspie Speech Products, Burlington, Massachusetts.
Jane Wozniak, Speech and Language Pathology, Private Therapy, Massachusetts.

Graduate Students
Lan Chen, Harold Cheyne, Laura Dilley, Heather Gunter, Andrew Howitt, Annika Karlsson-Imbrie, Roy Kim, Xiaomin Mou, Hale Ozsoy, Kelly Poort, Ariel Salomon, Janet Slifka, Jason Smith, Atiwong Suchato, Virgilio Villacorta,

Undergraduate Students
Sasha Devore, Emily Hanna, Stefan Hurwitz, Anna Khasin, Shuley Nakamura, Alice Suh, Desiree Syn, Helen Tsai, Jeremy Vogelmann,

Technical and Support Staff
Arlene E. Wint

1 Constraints and Strategies in Speech Production

1.1 Development of facilities

Physiological/acoustic recording and analysis. We have completed the implementation of a new data acquisition system for our EMMA movement transducer and have begun to run experiments with subjects. We recruited 12 subjects and recorded four of them in a motor equivalence experiment. We have also run a combined movement and EMG experiment, to begin providing data for the development of computational models of speech articulations by the vocal tract. We have begun the development of multi-channel, multi-bandwidth data analysis facilities and have extracted movement and acoustic data on the initial subjects. We have had constructed a lingua-palatal pressure recording system and are in the process of integrating the system into our overall experimental setup.

Two-dimensional vocal tract modeling. We refined our 2-D physiologically based biomechanical tongue model by fine-tuning individual muscles to produce agreement with x-ray data from the subject on whom the model is based. To complete this model, we incorporated jaw rotation and translation and lip opening and protrusion, using second order systems of differential equations. The complete model was then optimized to generate subject-specific formants (F1, F2, F3, and F4) with ±7.5% error for the vowels /i/, /e/, /a/, /o/, and /u/. We ran a large number of simulations with the model and used the resulting data to approximate (±5% error) the mapping between the muscle lengths and acoustic space (Forward Model). We also developed an algorithm for the control of the 2-D model of the vocal tract. The algorithm takes as input a desired movement direction in acoustic space and converts it into a set of optimal muscle length changes appropriate for control of the biomechanical model to produce the desired acoustic change (Inverse Model). In effect, the Inverse Model calculates the pseudoinverse of the Jacobian transformation relating acoustic trajectories to motor commands.

Three-dimensional vocal-tract modeling. To obtain a finite element model of the oral anatomy, two data structures were generated using the Visible Human data set as a reference. The first is a geometrical block composition of the shape of the tongue blade and floor of the mouth, in which the model is composed of large curvilinear hexahedral blocks that can be easily subdivided for finite element mesh generation. In the second structure, each muscle is represented independently as one or several blocks with curvilinear surfaces. Fiber directions inside these blocks are obtained as tangents of curvilinear coordinate lines. This muscle block representation is independent of the first data structure. During the finite element mesh generation, the muscle representations are used for an automatic computation of muscle fiber directions for each muscle throughout the finite element mesh.

Sensorimotor Adaptation studies. To enable us to run sensorimotor adaptation studies, we have written software to perturb the speech waveform produced by a speaker when uttering the phoneme /r/ in a carrier phrase. This software decreases the dip in the third formant frequency (F3) that characterizes /r/. When fed back to the ears of a speaker so that he/she can hear his/her own productions in near-real time, this perturbation is expected to cause the speaker to produce a compensatory increase in the F3 dip. Work is in progress to implement a system on a DSP board for performing such transformations in nearly real time. With this system, the perturbed version of a subject's own speech signal will be fed back to the subject and we will look for articulatory adjustments to compensate for the perturbations.

2. The Role of Hearing in Speech: Cochlear Implant Users

2.1 Rapid changes in speech production parameters in response to a change in hearing.

The speed of changes in speech production parameters was investigated in seven cochlear implant users, in response to switching the speech processors of their implants on and off a number of times in a single experimental session. The subjects repeated short utterances many times (in semi-random order). The switches between hearing (on) and non-hearing (off) states were introduced between utterances; the number of utterances between switches was varied to minimize subject anticipation of the switches. Two normal-hearing subjects performed the same paradigm, except that the on condition consisted of hearing their own speech fed back to them over a set of headphones and the "off" condition consisted of hearing loud noise that masked their speech. Using the times of on-off or off-on switches as line-up points for averaging, the following parameters were compared across the switches: median and symmetry of sibilant spectra, and F1, F2, duration, and SPL of vowels. The speakers' vowel SPL and duration had changed by the first utterance following the switch. Changes in contrast between phonemes were less prevalent, but just as immediate. The nature and speed of these changes indicate that they are part of the same mechanism, one that attempts to maintain intelligibility in the face of changing acoustic transmission conditions.

2.2 Speech perception, production and intelligibility improvements in vowel-pair contrasts among adults who receive cochlear implants.

This study investigated relations among speech perception, speech production and intelligibility in postlingually deaf adults who receive cochlear implants. Measures of the three variables for eight vowel pairs in acoustic space were gathered from eight post-lingually deaf adults pre- and post-implant. Improvements in a speaker's production, perception and intelligibility of a given vowel contrast tended to occur together. Subjects who produced vowel pairs with reduced contrast (measured by separation in the acoustic vowel space) pre-implant and who showed improvement in their perception of these contrasts (measured with a phoneme recognition test) post-implant were also found to have improved production contrasts. These enhanced production contrasts were associated with enhanced speaker intelligibility (tested with normal-hearing listeners). The results support the hypothesis that the implant users' improved speech perception was responsible, at least in part, for their improved speech production.

2.3 Changes in noise-masked word intelligibility of postlingually deaf adults after cochlear implantation.

This study examined changes in word intelligibility of postlingually deaf adults after they have had extensive experience with a cochlear implant. The speech of eight postlingually deaf adults was recorded before activation of the speech processors of their cochlear implants and at six months and one year after activation. Ten listeners with normal hearing completed a word identification task while listening to each implant user's speech in noise. The percent information transmitted by the speakers in their pre- and post-activation recordings was measured for 11 English consonants and eight vowels separately. An overall improvement in word intelligibility was observed: seven of the eight speakers showed improvement in vowel intelligibility and six speakers showed improvement in consonant intelligibility. However, the intelligibility of specific consonant and vowel features varied greatly across speakers.

2.4 Language-specific, hearing-related changes in vowel spaces: a preliminary study of English- and Spanish-speaking cochlear implant users (Done in collaboration with and partly funded by the Dept. of Otolaryngology, University of Miami Medical School.)

This study was designed to investigate the role of hearing in controlling vowel production of postlingually deafened cochlear implant (CI) users. We hypothesize that vowel production is influenced by competing demands of intelligibility for the listener and least effort for the speaker. Hearing should enable a CI user to produce vowels distinctly from one another; without hearing, the speaker may default somewhat to economy of effort, leading to reduced vowel contrasts. Furthermore, speakers may need to produce vowels more distinctly from one another in a language with a relatively “crowded” vowel space, such as American English, than in a language with relatively few vowels, such as Spanish. Thus, when switching between hearing and non-hearing states, English speakers may show this hypothesized tradeoff between vowel distinctness and least effort, while Spanish speakers may not. Spanish speakers, whose vowel regions have centers that are further apart, may, when hearing is available, place less weight on distinctness than English speakers, allowing for more influence of least effort. When hearing is not available to the Spanish speakers, they will continue to use a strategy dominated by least effort. Therefore, this experiment was to test the prediction that there will be a reduction of average vowel spacing (AVS – average inter-vowel distance in the F1-F2 plane) with interrupted hearing for English-speaking CI users, but less systematic change in AVS for Spanish CI users. Vowel productions of seven English- and seven Spanish-speaking CI users were recorded when their implant speech processors were on and when they were off. AVS was consistently larger for the English speakers with hearing than without hearing. The direction of AVS change was more variable for the Spanish speakers, within and between subjects. Considered as a group, the English speakers reduced AVS and the Spanish speakers did not. Thus the prediction is upheld, and overall, the results illustrate a possible trade-off between the distinctness of acoustic goals and economy of effort.

2.5 Effects of cochlear implants on the production, perception, and intelligibility of the liquids /r/ and /l/

Speech production, perceptual testing and speech intelligibility data for /r/ and /l/ were obtained from eight postlingually deaf adults, pre- and post cochlear implant (CI). Formant transition analysis for the CI speakers and two speakers with normal hearing indicated that /r/ and /l/ could be reliably differentiated by the extent of the F3 transition and the distance in Hz between F2 and F3 at the consonant-vowel boundary. Speakers who had a limited contrast between /r/ and /l/ pre-cochlear implant and who showed improvement in their perception of these consonants with prosthetic hearing were found to demonstrate greatly improved production of /r/ and /l/ six months post-CI. The speech production changes noted in the acoustic analyses were corroborated by intelligibility improvements in the post-CI speech, as measured with a panel of normal-hearing listeners.

2.6 Changes in voice production with hearing loss

This study explored the effect of postlingually acquired deafness on voice parameters using longitudinal voice production data from subjects who have cochlear implants (CI) or Neurofibromatosis II (NF2). CI subjects were recorded pre-implant and up to two years post-implant. The NF2 subjects had unilateral profound deafness due to tumor surgery prior to the beginning of the study, with significant residual hearing in their other ear. During the study, three of the NF2 subjects became profoundly deaf due to surgery to remove acoustic neuromas on their previously unaffected sides, and two of these subjects received auditory brainstem implants (ABI). Voice analyses included SPL, F0, peak and average airflow, and intra-oral pressure measures. Results from the cochlear-implant subjects and NF2 subjects were compared with normative values. The most significant changes due to hearing status were found for the three speakers with NF2 who had become profoundly deaf. When deaf subjects were asked to produce loud voice speech, results were widely variable, suggesting that this was a difficult task without access to auditory feedback.

3 Models for Lexical Access

3.1 Enhancement and gestural overlap as sources of variability in speech production

When an utterance is planned by a speaker, it is generally assumed that the sequence of words is stored in a planning stage in the form of sequences of phonological segments, each of which consists of an array of binary distinctive features. The speaker's task is to convert this discrete linguistic description into a set of continuous motor commands or gestures of the respiratory, laryngeal, and articulatory structures. The movements of these structures ultimately result in the generation of sound in the vocal tract and the radiation of this sound from the mouth and/or nose.

The distinctive features in the planning stage are drawn from a universal inventory of features; a given language selects a subset of these features to define the contrasts in the language. Each feature provides a sketch of the articulatory actions that are used to implement the feature. The features are of two kinds: articulator-free features that specify a general articulatory actions, such as making a consonantal constriction in the vocal tract, and articulator-bound features that specify which articulators are to be manipulated and how they are to be positioned. Associated with each feature there are basic acoustic correlates that define the perceptual contrast for that feature.

In the process of transforming the (invariant) linguistically specified planning stage to a sound output, two kinds of variability can be introduced. One of these, which we call enhancement, occurs when, for a given feature in a particular phonetic environment, the acoustic and perceptual contrast provided by the basic acoustic correlates for the feature is weak. Additional articulatory actions are then recruited to enhance this contrast. This enhancement is intended to strengthen the defining acoustic attribute for the feature, and it may also introduce additional acoustic cues that a listener can use to identify the feature. Thus the inventory of acoustic cues that are available to identify a feature may depend on the context in which the feature occurs, and consequently there is variability in the mapping from sound to underlying features.

A second source of variability in the acoustic speech signal arises because the gestures used to implement the features for adjacent segments may overlap. This overlap can obscure some of the acoustic cues for the features. An example is the sequence “up to”, in which the closure for /t/ overlaps with the release for /p/, so that there is no direct evidence in the sound for these two articulatory events. It appears, however, that if an enhancing gesture is introduced for a particular feature, acoustic evidence for this gesture is never eliminated as a consequence of gestural overlap. However, elimination of a basic acoustic correlate for a feature through gestural overlap can occur if evidence for the feature is available through enhancing gestures. The general principle, then, is that acoustic evidence for a distinctive feature is never eliminated, but this evidence may appear in a form different from the form predicted from the basic acoustic correlates of the feature.

As we develop and implement a model for human lexical access from running speech, we expect to test the validity of this view of variability in speech.

3.2 Identification of vowel landmarks

An initial speech processing step in our proposed model for lexical access is the detection of several classes of acoustic landmarks: peaks in low-frequency amplitude for vowels, minima in low-frequency amplitude (without abrupt discontinuities) for glides, and spectral discontinuities at closures and releases for consonants. These landmarks indicate the presence of underlying vowel, consonant, and glide segments, and specify regions where acoustic analysis should be concentrated to identify articulator-bound distinctive features for these segments. Initial versions of landmark detectors for consonants and glides have been developed and evaluated some years ago, and a landmark detector for vowels has now been completed. Analysis of vowels in a large database of sentences has shown that about 6 percent of vowels do not exhibit a peak in low-frequency amplitude, indicating that a vowel detector based on this simple acoustic property will fail to detect these vowels. The current vowel detector is based on a peak-picking technique, using a recursive convex hull algorithm. Three acoustic cues (peak-to-dip depth, duration, and level) were combined using a multilayer perceptron with two hidden units. The perceptron was trained by matching nuclei derived from a labeled database. The final error rate for the vowel detector was about 12 percent, with about 3.5 percent insertions and 8.5 percent deletions (i.e., slightly greater than the 6 percent of vowels without peaks). Most of the deletions were from abutting vowels, with the second vowel being reduced. Further more refined analysis of other acoustic parameters will be necessary to detect the presence of these deleted vowels and to detect that an erroneous insertion is not a separate vowel.

4 Prosodic Influences on Phonetic Variation

In addition to work on a model of human analysis of speech signals for lexical access, our progress in the past year included a) development of a conceptual model of phonological encoding, based on observations of patterns of surface phonetic variation in word form as well as speech error data, b) determination of phonological differences between monosyllabic function words and other word categories of American English which may be related to their tendency toward severe phonetic modification in context, and c) demonstration of progressive lengthening from the onset to the coda of a phrase-final syllable. These lines of research are consistent with the hypothesis that articulatory patterns preserve (and in some cases enhance) acoustic cues to the underlying distinctive contrasts that define word forms. Results support the hypothesis that the seemingly unsystematic variation which we observe in the surface phonetic forms of words across contexts is systematically governed, at least in part, by the same prosodic structure that specifies the suprasegmental aspects of an utterance, such as its intonation and timing.

a) The conceptual model of phonological encoding (in collaboration with P. Keating of UCLA) contrasts with existing models in its focus on i) the top-down construction of a prosodic framework for the utterance, ii) retrieval of various aspects of word-form information, such as metrical structure (number of syllables and stress pattern) and segmental information, only as they are needed for computation of the prosodic representation, and iii) the need to account for the systematic variation in word forms across different segmental and prosodic contexts. In this model, construction of the prosodic framework initially maps the surface syntactic structure of the sentence onto constituents of the prosodic hierarchy (e.g., the utterance, intonational phrase, phonological phrase, prosodic word etc.) This initial representation is subsequently restructured on the basis of non-grammatical information, such as the length of each constituent (provided by the metrical information retrieved from the lexicon for each word) and the speaking rate. Only after this restructuring is the segmental information about the words retrieved. Finally, the segmental and prosodic specifications of the utterance are integrated into parameter values that govern the timing and amplitude of the distinctive-feature-specified articulatory gestures.

This model provides an account of sound-level phenomena which involve more than one word, such as cross-word-boundary assimilations, exchange errors, effects of constituent length on phrasing etc., as well as for the increasingly-well-documented effect of hierarchical prosodic constituent structure on phonetic variation. It also accounts for the process of translating the abstract categories of phonemes and distinctive features (which distinguish words in the lexicon) to quantitative articulatory specifications in terms of continuous-valued timing and constriction-degree parameters (which control the vocal tract). Ongoing testing of the model with speech error data and acoustic analysis of the nature of surface phonetic variation continues in our laboratory.

b) The special phonological characteristics of monosyllabic function words in American English were analyzed using the Brown corpus of more than 1 million words of text which is labeled for the syntactic class of each word in context. Content words (i.e. members of the open syntactic classes of Nouns, Verbs, Adjectives, and Adverbs formed with the –ly suffix) were compared with function words (i.e. members of the closed syntactic classes such as Articles, Conjunctions, Prepositions, Particles, Pronouns, etc.). The phonological structure of function words differs from that of content words in that function words are more likely (1) to be monosyllabic, (2) to begin with a vowel, and, if they begin with a consonant, (3) to begin with a more sonorant consonant than do content words. These findings hold for both the lexicon (where each word is represented only once) and in the corpus (where the high frequency of the function words exacerbated the content word/function word differences). We suggest that the more impoverished syllable structure and more sonorant consonantal inventory of function words may be related to their renowned susceptibility to extreme phonetic modification and reduction in connected speech, as in e.g., 'Whyncha' for 'Why don't you'.

c) Progressive lengthening from the onset to the coda of a phrase-final syllable was found for American English, in spoken sentences which contrast the placement of utterance-internal phrase boundaries, e.g. after 'Maine' in 'Please say Maine, or Duke and Rice will play', but after 'Duke' in 'Please say Maine or Duke, and Rice will play'. Direct comparison of durations in pre-boundary vs. non-preboundary conditions showed that the onset consonant increased only negligibly in duration, the vowel increased somewhat and the final consonant increased the most. This progressive lengthening in the final syllable of an utterance-internal intonational phrase extends our understanding of the effects of prosodic structure on the realization of the segmental contrasts which specify the words of an utterance.

5 Studies of Normal Speech Production

5.1 Speaker characteristics, longitudinal change, and acoustic cues for distinctive features

A database of citation-form consonant-vowel-consonant utterances was collected some years ago from three speakers. The speakers recorded the utterances on two occasions 30 years apart. An initial motivation for examining these recordings was to observe some of the ways in which the acoustic characteristics of a speaker’s utterances can change over this 30-year period. As the analysis of the recordings has proceeded, two other uses of the material have emerged. The recordings can help to characterize what aspects of a speaker’s speech production remain constant over the years; i.e., what aspects characterize a particular speaker in spite of possible significant changes in other aspects. A second use of the material is to gain insight into the acoustic attributes associated with particular phonemic features that remain in spite of speaker differences and in spite of changes in characteristics of particular speakers with aging.

A case study of the data has examined the age-related changes that occur in the spectrum of the glottal source during phonation and the changes in the spectra of the noise bursts that occur at the release of obstruent stop consonants. The production mechanisms that influence the glottal source spectrum would appear to be the result of changes in laryngeal physiology, whereas it is expected that turbulence noise at a consonant release should be determined primarily by movements of supralaryngeal structures. Thus one might expect age-related changes in these two attributes of speech production to be independent. An alternative view, however, is that a listener (and a speaker) assesses the acoustic attributes for a particular stop-consonant place of articulation in relation to the acoustic spectrum of an adjacent vowel. In such a view, a speaker who exhibits an age-related change in the glottal source spectrum during vowels might be expected to modify the production of consonant releases in order to produce a spectrum that bears an appropriate relation to vowels generated with the modified glottal source. Thus the data can be used to determine whether the place-of-articulation acoustic cues provided by a noise burst should be assessed in relation to the acoustics of the glottal source for the adjacent vowel. Such a view would predict a close relation between age-related changes in the glottal source for vowels and the noise source for stop consonants.

Of the three speakers who generated the database, two showed a significant increase in the high-frequency “tilt” of the glottal source spectrum over the 30-year period, and one did not. The two speakers who show a change in the glottal source spectrum exhibit corresponding changes in the spectrum of the turbulence noise spectrum for the burst for alveolar and labial stop consonants. This result suggests that those speakers who showed an increase in the slope of their glottal source spectrum adjusted the spectra of their stop-consonant bursts in a direction that compensated at least partially for the changed vowel spectra. That is, the speakers adapted their production of the stop consonants to their age-related changes in the glottal source for vowels.

5.2 Respiration and prosody in speech production

The aim of this research on respiration in speech production has been to quantify some of the constraints that the respiratory system imposes on the acoustic properties of the beginnings and ends of phrases and in the vicinity of pauses within a phrase. Data on pressures, flows, and lung volume during a number of utterances by several speakers have been obtained in collaboration with the Voice Laboratory at the Massachusetts Eye and Ear Infirmary in Boston. These data, combined with calibration runs for the speakers, have provided estimates of the contributions of the active inspiratory and expiratory forces and of the relaxation forces of the chest wall in determining the time variation of pleural pressure as lung volume decreases during an utterance. The data from pressures and flows also permit estimates of the cross-sectional area of the glottal opening. The data show that an utterance is usually initiated near the end of an inspiration, while the inspiratory forces are still active, but there is a rapid transition from inspiratory effort to forces resulting from chest wall relaxation. Near the termination of an utterance, the expiratory forces are active, but there is usually a gradual drop in the lung pressure, and the estimated glottal area can show as much as a threefold increase relative to the average area for normal phonation. As the pressure drops near the end of phonation, irregular glottal vibration is often observed. If a pause is inserted during a longer phrase, it is usually implemented by a drop in pressure to 2-3 cm H_2O, with this drop occupying a time interval of about 300 milliseconds. Lung volume changes during such a pause are minimal. The changes in acoustic amplitude and glottal source spectrum in the vicinity of these boundary events can be explained in terms of these observations of respiratory actions. One application of this research is to develop improved procedures for the machine synthesis of utterance beginnings, endings, and pauses.

Publications

Journal Articles, Published

Chen, M.Y., K.N. Stevens, H.-K. Kuo and H. Chen. "Contributions of the Study of Disordered Speech to Speech Production Models." J. Phonetics 28: 303-312 (2000).

Matthies, M., Perrier, P., Perkell, J. and Zandipour, M. Variation in Coarticulation with Changes in Clarity and Rate." J. Speech, Lang. Hear. Res. 44: 340-353 (2001).

Perkell, J., F. Guenther, H. Lane, M. Matthies, P.Perrier, J.Vick, R. Wilhelms-Tricarico, and M. Zandipour. "A Theory of Speech Motor Control and Supporting Data from Speakers with Normal Hearing and with Profound Hearing Loss." J. Phonetics 28: 233-372 (2000).

Stevens, K.N. "Diverse Acoustic Cues at Consonantal Landmarks." Phonetica 57: 139-151 (2000).

Stevens, K.N. "Recognition of Continuous Speech Requires Top-Down Processing." Brain and Behavioral Sciences 23: 348 (2000).

Stevens, K.N. "The Properties of the Vocal-Tract Walls Help to Shape Several Phonetic Distinctions in Language." In Travaux du Cercle Linguistique de Copenhague, Vol. XXXI: 285-297 (2001).

Journal Articles, Accepted for Publication

Turk, A.E., and S. Shattuck-Hufnagel. "Word-Boundary-Related Duration Patterns in English." J.Phonetics, Forthcoming.

Journal Articles, Submitted for Publication

Chen, H. and K.N. Stevens. "An Acoustical Study of the Fricative /s/ in the Speech of Individuals with Dysarthria," submitted to J. Speech, Language, and Hearing Research.

Gould, J., H. Lane, J. Perkell, J. Vick, M. Matthies, and M. Zandipour. "Changes in the Intelligibility of Postlingually Deaf Adults after Cochlear Implantation," submitted to Ear and Hearing.

Hanson, H.M., K.N. Stevens, H.-K. Kuo, M.Y. Chen, and J. Slifka. "Towards Models of Phonation," submitted to J. Phonetics.

Kwong, K.W. and K.N. Stevens. "On the Voiced-Voiceless Distinction for Writer/Rider," submitted to J. Phonetics.

Lane, H., M.L. Matthies, J.S. Perkell, J. Vick, and M. Zandipour. "The Effects of Changes in Hearing Status in Cochlear Implant Users on the Acoustic Vowel Space and Coarticulation," submitted to J. Speech. Lang. Hear. Res.

Matthies, M., P. Perrier, J. Perkell, and M. Zandipour. "Variation in Speech Movement Kinematics and Temporal Patterns of Coarticulation with Changes in Clarity and Rate," submitted to J. Speech, Lang. Hear. Res.

Matthies, M., J. Vick, J. Perkell, H. Lane, M. Zandipour, and J. Gould. "Effects of Cochlear Implants on the Speech Production, Perception, and Intelligibility of the Liquids /r/ and /l/," submitted to J. Speech, Lang. Hear. Res.

Perkell, J., W. Numa, J. Vick, H. Lane, T. Balkany, and J. Gould. "Language-Specific, Hearing-Related Changes in Vowel Spaces: A Preliminary Study of English- and Spanish-Speaking Cochlear Implant Users," submitted to Ear and Hearing.

Perkell, J., M. Zandipour, J. Vick, M. Matthies, H. Lane, and J. Gould. "Rapid Changes in Speech Production Parameters in Response to a Change in Hearing, " submitted to J. Phonetics.

Vick, J., H. Lane, J. Perkell, M. Matthies, J. Gould, and M. Zandipour. "Speech Perception, Production and Intelligibility Improvements in Vowel-Pair Contrasts in Adults who Receive Cochlear Implants," submitted to J. Speech, Lang. Hear. Res.

Perkell, J.S., and F.H. Guenther. "Speech Motor Control: Acoustic Goals, Saturation Effects, Auditory Feedback and Internal Models," submitted to J. Phonetics.

Stevens, K.N. "Toward a Model for Lexical Access Based on Acoustic Landmarks and Distinctive Features," submitted to J. Acoust. Soc. Am.

Book/Chapters in Books

Shattuck-Hufnagel, S. "Phrase-level Phonology in Speech Production Planning: Evidence for the Role of Prosodic Structure." In *Prosody: Theory and Experiment: Studies presented to Gosta Bruce*. Ed. Merle Horne. Stockholm:Kluwer, 2000.

Keyser, S.J. and K.N. Stevens. "Enhancement Revisited." In *Ken Hale: A Life in Language.* Ed. M. Kenstowic. Cambridge MA: MIT Press, 271-291, 2001.

Meeting Papers, Presented

Matthies, M., J. Perkell, J. Vick, and M. Zandipour, "Cochlear-Implant Effects on Sibilants." Paper presented at the Acoustical Society of America Meeting, Newport Beach, California, December 3-8, 2000.

Perkell, J.S., and F.H. Guenther, "A Model of Speech Motor Control and Supporting Data: Influences of Quantal Effects." Invited Paper presented at the Acoustical Society of America Meeting, Newport Beach, California, December 3-8, 2000.

Shattuck-Hufnagel, S., “Right Boundary Strengthening in American English.” Paper presented at the Acoustical Society of America Meeting, Newport Beach, California, December 3-8, 2000.

Stevens, K.N., and S.J. Keyser, “Retrieving Distinctive Features and Segments from Variable Acoustic Cues.” Invited Paper presented at the Acoustical Society of America Meeting, Newport Beach, California, December 3-8, 2000.

Vick, J., H. Lane, J. Perkell, M. Matthies, J. Gould, and M. Zandipour, “Speech Perception, Production, and Intelligibility Improvements in Vowel-Pair Contrasts among Adults who Receive Cochlear Implants.” Paper presented at the Acoustical Society of America Meeting, Newport Beach, California, December 3-8, 2000.

Meeting Papers, Published

Chen, M.Y. “Nasal Detection Module for a Knowledge-Based Speech Recognition System.” Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP2000) Vol. IV, pp. 636-639, Beijing, China, 2000.

Howitt, A.W. “Vowel Landmark Detection.” Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP2000) Vol. IV, pp. 628-631, Beijing, China, 2000.

Perkell, J., M. Zandipour, J. Vick, M. Matthies, H. Lane, F. Guenther, and J. Gould. Rapid Changes in Speech Production Parameters in Response to a Change in Hearing. Proceedings of the 5th Seminar on Speech Production & CREST Workshop on Models of Speech Production , May 1-4, 2000.

Shattuck-Hufnagel, S. and N. Veilleux. “The Special Phonological Characteristics of Function Words in American English.” Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP2000) Vol.1, pp. 540-543, Beijing, China, 2000.

Stevens, K.N. “From Acoustic Cues to Segments, Features and Words.” Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP2000) Vol. I, pp. A1-A8, Beijing, China, 2000.

Turk, A.E., and S. Shattuck-Hufnagel. “Duration as a Cue to Syllable Affiliation.” Proceedings of the Conference on the Phonological Word, Berlin, October 1997.

Reports

Theses

Howitt, A.W. *Automatic Syllable Detection for Vowel Landmarks*. Ph.D. thesis, Department of Electrical Engineering and Computer Science, MIT, 2000.

Kim, Roy K.-H. *Implementing the Matcher in the Lexical Access System with Uncertainty in Data*. M.Eng. thesis, Department of Electrical Engineering and Computer Science, MIT, 2000.

Ozsoy, H.H. *An Acoustic Study of Coarticulation Patterns in Turkish and English*. M.Eng. thesis, Department of Electrical Engineering and Computer Science, MIT, 2000.

Poort, K.L. *Acoustic and Perceptual Assessment of Stop Consonants Produced by Normal and Dysarthric Speakers*. Ph.D. thesis, Harvard-MIT Division of Health Sciences and Technology, MIT, 2000.

Slifka, J. *Respiratory Constraints on Speech Production at Prosodic Boundaries*. Ph.D. thesis, Harvard-MIT Division of Health Sciences and Technology, MIT, 2000.