Chapter 1. Speech Communication

Academic and Research Staff

Professor Kenneth N. Stevens, Professor Jonathan Allen, Professor Morris Halle, Professor Samuel J. Keyser, Dr. Corine A. Bickley, Dr. Marilyn Chen, Dr. Joseph S. Perkell, Dr. Stefanie Shattuck-Hufnagel, Dr. Reiner Wilhelms-Tricarico, Seth M. Hall, Jennell C. Vick, Majid Zandipour

Visiting Scientists and Research Affiliates

Dr. Ashraf Alkhairy,[1] Dr. Suzanne E. Boyce,[2] Sherrie Brown,[3] Dr. Carol Y. Espy-Wilson,[4] Dr. Krishna Govindarajan,[5] Dr. David Gow,[6] Dr. Frank Guenther,[7] Dr. Mark Hasegawa-Johnson,[8] Dr. Helen M. Hanson,[9] Dr. Robert E. Hillman,[10] Dr. Caroline Huang,[11] Dr. John Ingram,[12] Dr. Harlan Lane,[13] Dr. John I. Makhoul,[14] Dr. Sharon Y. Manuel,[15] Dr. Melanie L. Matthies,[16] Dr. Richard McGowan,[17] Dr. Alice Turk,[18] Dr. Lorin Wilde,[19] Jane Wozniak[20]

Graduate Students

Helen Chen, Harold A. Cheyne, Jeung-Yoon Choi, Laura C. Dilley, Michael P. Harms, Dameon C. Harrell, Andrew W. Howitt, Hong-Kwang J. Kuo, Aaron T. Maldonado, Kelly L. Poort, Adrienne M. Prahler, Andrew I. Russell, Janet L. Slifka, Jason L. Smith, Felice T. Sun, Yong Zhang

Undergraduate Students

Emily Hanna, Stefan Hurwitz, Mariya Ishutkina, Genevieve Lada, Teresa Lai, Hale Ozsoy, Dawn Perlner, Delsey Sherrill, Svetla Tzenova, Jeremy Vogelmann, Sophia Yuditskaya

Technical and Support Staff

Arlene E. Wint

---

1 KACST, Riyadh, Saudi Arabia.

2 Department of Communication Disorders, University of Cincinnati, Cincinnati, Ohio.

3 Department of Neuropsychology, Massachusetts General Hospital, Boston, Massachusetts.

4 Department of Electrical Engineering, Boston University, Boston, Massachusetts.

5 SpeechWorks International, Boston, Massachusetts.

6 Department of Psychology, Salem State College, Salem, Massachusetts.

7 Department of Cognitive and Neural Systems, Boston University, Boston, Massachusetts.

8 Department of Electrical Engineering, University of California at Los Angeles, Los Angeles, California.

9 Sensimetrics Corporation, Somerville, Massachusetts.

10 Massachusetts Eye and Ear Infirmary, Boston, Massachusetts.

11 SpeechWorks International, Boston, Massachusetts.

12 Department of English, University of Queensland, St. Lucia, Australia.

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

14 GTE Internetworking, BBN Technologies, Cambridge, Massachusetts.

15 Department of Speech-Language Pathology and Audiology, Northeastern University, Boston, Massachusetts.

16 Department of Communication Disorders, Boston University, Boston, Massachusetts.

17 Sensimetrics Corporation, Somerville, Massachusetts.

18 Department of Linguistics, University of Edinburgh, Edinburgh, United Kingdom.

19 Lernout and Hauspie Speech Products, Burlington, Massachusetts.

20 Concord Area Special Education (CASE) Collaboratives, Concord, Massachusetts.

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 DC02525-01A1, 1-F32-DC00194, 1-F32-DC00205, T32-DC00038
National Science Foundation
Grants IRI-9314967,[21] INT-9421146,[22] and INT-9615380[23]

1.1 Studies of Normal Speech Production

1.1.1 Constraints and Strategies In Speech Production: Motor Equivalence in the Production of /ʃ/

To explore the idea that speech motor goals are acoustic targets, upper lip protrusion and tongue blade fronting were examined in the sibilant /ʃ/ for evidence of motor equivalence in eight speakers of American English. Positive correlations across multiple repetitions of /ʃ/ (i.e., motor equivalence) would occur if the upper lip compensated with more protrusion when the tongue blade was further forward and vice versa. This motor equivalence would serve to maintain an adequate front cavity volume, thus enhancing the acoustic stability of /ʃ/ and assuring good acoustic separation from /s/. It was hypothesized that motor equivalence would be found among acoustically less canonical tokens, as elicited in a "clear+fast" speaking condition (used in addition to a normal condition). Acoustic spectral analyses indicated excellent separation between the two sibilants, even in the clear+fast condition. Four of the eight subjects showed some significant positive correlations of tongue-blade and upper lip displacements. When there was a difference between conditions in how canonical the tokens were, the motor equivalent tokens were less canonical. There were no significant negative correlations. There were relatively few motor equivalence findings, perhaps because some speakers use saturation (quantal) effects, which could also enhance acoustic stability.

1.1.2 Respiration and Prosody in Speech Production

The production of speech requires control of three groups of structures: (1) the articulators that form the vocal tract above the larynx, (2) the larynx, and (3) the respiratory system. The respiratory system regulates the air in the lungs to maintain breathing for life purposes and to provide energy for sound sources during speaking. This dual purpose and the dynamics of the respiratory system impose constraints on (1) how utterances are initiated and terminated and (2) the realization of prominences and pauses in the utterances.

We have begun a series of experiments that examine the lung volume, subglottal pressure, and airflow used by normal speakers when they produce various kinds of sentence material, together with several acoustic measures. The data obtained from these experiments will be used to develop a model of the respiratory system which can be used as a component of a speech synthesizer. Initial experiments have observed the time course of the rise in subglottal pressure at the beginning of an utterance and the fall in pressure near the end of an utterance and have examined the influence of these respiratory changes in the characteristics of the acoustic source at the glottis.

Preliminary observations of pauses within a breath in an utterance show small decreases in subglottal pressure but general avoidance of loss of respiratory volume through such actions as the formation of a glottal or supraglottal closure during the pause. The measurements of respiratory parameters are being carried out in collaboration with the Voice Laboratory at the Massachusetts Eye and Ear Infirmary, under the direction of Dr. Robert Hillman.

21 Under subcontract to Stanford Research Institute (SRI) Project ECU 5652.
22 U.S.-India Cooperative Science Program.
23 U.S.-France Cooperative Research Program.

1.2 Speech Research Relating to Special Populations

1.2.1 The Speech of Cochlear Implant Patients

In order to characterize the speech production of postlingually deafened adults before and after they receive cochlear implants (CIs), our research aims to describe the disordered speech that follows deafening in adulthood and the changes in speech that follow the restoration of some hearing with CIs. Moreover, we are using the findings from implant users and normally hearing adults to help refine a theoretical framework and a derived descriptive model of the role of hearing in adult speech production.

Longitudinal Studies

We have made three baseline pre-implant recordings on each of five additional subjects (three implant users and two with normal hearing). Nine speakers returned for post-implant recordings one week, one month, and three months following processor activation. Six of our nine implant subjects have returned for their one-year visits, and the remaining three will complete six-month recordings in April 1999. Thirty-three recordings in all have been made this year with CI subjects. We recruited two normal-hearing subjects to be used as controls and have made a total of six recordings of their speech. Over 65% of our recordings have been digitized, and more than half of the data extraction and analysis has been completed.

Short-term Stimulus Modification Studies

The speed of changes in speech production parameters was investigated in two cochlear implant users. The speech processors of their implants were switched on and off a number of times while subjects read short utterances with repetitions in semi-random order. The transitions between hearing (on) and non-hearing (off) states were timed to occur between utterances, and the number of utterances between transitions was varied to minimize subject anticipation of the change. Using the times of on-off or off-on transitions as line-up points, values of median symmetry and symmetry of sibilant spectra, together with vowel *F*1, *F*2, duration and SPL, were each averaged over repetitions of each utterance and compared across the transitions. Vowel SPL and duration changed by the first utterance following the transition, indicating that the subjects were using subtle aspects of the processor output to detect its state even in the absence of overt sound input. Changes in spectral median and symmetry and *F*1 and *F*2 were less prevalent, but just as immediate. The paradigm is being refined and additional subjects are being run.

Perceptual Studies

At the time of each speech recording, subjects are asked to discriminate eleven consonants and eight vowels from the natural speech of a same-gender normally hearing speaker. Initial analyses have been performed on these data using measures of information transmitted for various features. Voicing and manner features are more robustly transmitted than place features. We aim to assess whether implant users' evolving abilities to perceive and produce various features covary.

Intelligibility Testing

Preparations have been completed for testing the intelligibility of implant users' speech before and at various intervals after the initial activation of their speech processors. Three repetitions of each of eight vowels and eleven consonants were extracted from speaker's two pre-implant sessions and last two post-activation sessions and assembled in random order. In order to assess changes in intelligibility from pre- to post-implantation and over time, it will be necessary with some of the speakers who are highly intelligible to mix noise with their recordings before presentation to listeners. In order to determine the appropriate noise level, a preliminary experiment is planned with three different S/N ratios. The resulting recording will be presented in group listening sessions for multiple choice identification of the target tokens, with carefully selected written foils designed to reveal featural errors.

Acoustic Correlates of Speech Posture

Acoustic and aerodynamic measures were collected from seven adventitiously deafened cochlear implant users, three speakers who had severe reduction in hearing following surgery for Neurfibromatosis-2, and one hard-of-hearing speaker reading the Rainbow Passage and an English vowel inventory before and after intervention. All except one of the postlingually deaf speakers who received prosthetic hearing reduced speech sound level. Three of these significantly increased a measure of inferred glottal aperture, and their session means for these two parameters were inversely correlated longitudinally.

All except one of the speakers terminated respiratory limbs closer to functional residual capacity, once prosthetic hearing was supplied. Finally, the implant users' average values of air expenditure moved toward normative values with prosthetic hearing. These results are attributed to the mediation of changes in respiratory and glottal posture aimed at reducing speech sound level and economizing effort.

Coarticulation

Less coarticulation is hypothetically a feature of clear speech, and our dual process theory predicts that deafened speakers will engage in clear speech. Therefore, we predict that deafened adults will show less coarticulation before and more coarticulation after their implant speech processors have been turned on. We have begun to test this prediction by extracting values of second formant frequency from readings of the vowel inventory in /bVt/ and /dVt/ contexts by seven implant users. We have developed an algorithm to (1) assemble all the token trajectories for a given syllable; (2) assign to each trajectory a type classification; and (3) as a function of the trajectory and its type, assign a tentative *F*2 target. Data extraction has been completed for one hearing control speaker and is in progress for CI subjects.

Differential Expansion of Vowel Space

Several investigators have reported that the vowel space is reduced in size after prolonged deafness and expands with the provision of prosthetic hearing. We hypothesize that the vowel space will expand less among Spanish-speaking implant users than among those who speak English because Spanish has a smaller number of vowel phonemes and thus can achieve a given degree of separation among perceptual targets with a smaller space than can English. We plan to record ten matched pairs of Spanish- and English-speaking postlingually-deafened implant patients reading utterances containing the vowel inventories of their languages with their processors turned off and on. Three speakers have been recorded to date and data analysis is beginning.

1.2.2 Acoustic Characteristics Of Vowels Produced by Dysarthric Speakers

As part of a continuing study of the acoustic characteristics of dysarthric speakers, we have been examining the production vowels in isolated monosyllabic words. Twenty-nine words produced by six speakers with dysarthria (three males, three females) and four normal speakers (two males, two females) are being analyzed. General information on the dysarthric speakers, including the type of dysarthria, has been given previously in a doctoral dissertation by H.-P. Chang.[24] The acoustic measurements include formant movements and vowel durations and assessment of spectral characteristics that can be interpreted in terms of attributes of the glottal source.

In general, the vowels of the dysarthric speakers have greater durations than those of the normals, and the range of variation of the first two formant frequencies (*F*1 and *F*2) within the vowels of individual words is greater than for normals. There is greater fluctuation of *F*1 and *F*2 within the vowels for the dysarthric speakers, giving evidence for tremor in vowel production. Speakers with greater fluctuation tended to have lower overall intelligibility. Similar fluctuation is also observed in the frequency of glottal vibration. Some dysarthric speakers show glottal waveform characteristics outside the range observed for normals. For all the dysarthric speakers, there is evidence of more noise in the glottal source, based on observations of the vowel spectra and of vowel waveforms that are bandpassed in the range of the third formant frequency. The first two formant frequencies for individual dysarthric speakers tend to be shifted relative to the frequencies for normals, suggesting different basic postural positions for the tongue body. The shift is different for different speakers, and there is some indication that this postural shift depends on the type of dysarthria (spastic, ataxic, or athetotic), but the data are not sufficient to make a firm conclusion.

24 H.-P. Chang, *Speech Input for Dysartric Computer Users*, Ph.D. diss. Department of Mechanical Engineering, MIT, 1995.

1.3 Speech Production Planning and Prosody: Conversational Speech

As interest grows in the analysis, synthesis and automatic recognition of spontaneously spoken utterances, the characteristics that distinguish fluent conversational speech from laboratory speech and single-syllable utterances have become the focus of considerable interest. Our research this year has focused on three aspects of conversational speech which have significant effects on the acoustic-phonetic signal and which are not yet fully understood: the phrase-level prosody of spoken utterances, the nature of phonetic modification in spontaneous speech, and patterns of speech errors and disfluencies.

1.3.1 Effects of Prosodic Structure on the Acoustic-phonetic Shape of the Speech Signal

Our earlier work provided evidence that the hierarchy of prosodic constituents, such as two different levels of intonational phrase, predicts speaker behavior in such areas as early location of pitch accent within the word at the onset of an intonational phrase (as in "JAPanese FOOD" versus "He is japaNESE"), glottalization of phrase-initial and pitch accented vowels, and duration lengthening at the ends of prosodic constituents that increases monotonically with the location of the constituent in the prosodic hierarchy.

We have extended the work on duration, asking how far leftward in the phrase-final word the boundary-related lengthening is found, and whether it varies systematically across the sub-syllabic constituents of onset, nucleus and coda. In collaborative work with Alice Turk of the University of Edinburgh, we found that (1) the majority of lengthening occurs in the phrase-final syllable, whether that syllable is lexically-stressed or reduced; (2) some degree of lengthening is found leftward to the main-stress vowel of words like "Fairbanks," but not the reduced-syllable vowel of "Detroit"; and (3) duration lengthening in the phrase-final syllable increases monotonically from onset to coda. These results suggest that both higher-level prosodic structure (such as distinctions among intonational phrase types) and lower-level prosodic structure (such as lexical stress patterns and syllable substructure) are actively used by speakers in planning utterances for production and will also contribute to the naturalness of synthesized speech when incorporated into algorithms for computing systematic effects of phrase-level structure on segment duration.

Another line of investigation concerns the perception of rhythm. Rhythm is an aspect of spoken prosody which has not received extensive attention in quantitative studies. Our first task was to develop and test a reliable labeling algorithm for regions of rhythmic regularity in spoken utterances. A subsequent series of perceptual experiments has shown that two full-vowel syllables produced with a large $F0$ change between the syllables can be ambiguous with respect to the location of prominence. For example, the perceived pattern of prominent versus less prominent syllables in a sequence of monosyllabic full-vowel words like "I'm all right now," produced on a H-L-H-L intonation pattern, is more likely to be perceived as "I'M all RIGHT now" if preceded by a trochaic word (like "MAYbe") than if preceded by an iambic word (like "perHAPS"). Similarly, a string of full-vowel syllables like "cow town ship side way word list," produced on a H-L-H-L-H-L intonation contour, is heard as "cowtown shipside wayward." However, if the initial "cow" and final "ward" are removed, the same spoken string, now on a L-H-L-H contour, is reanalyzed as "township sideway"; moreover, for some listeners, the perceived prominence is now reversed, so that "town," formerly perceived as unstressed, is now perceived as stressed, in agreement with the known lexical stress pattern of the newly reanalyzed word. These results suggest that listeners must interpret the prominence patterns in at least some ambiguous strings of syllables, rather than registering them directly from the signal $F0$ characteristics.

1.3.2 Phonetic Modification

A second aspect of fluently-spoken utterances which is currently receiving more attention than in the past is the pattern of changes and variation that words, syllables and sounds undergo when they occur in longer phrases. The effects of adjacent words and sounds can be substantial, as can the effect of prosodic structure. For example, it has often been claimed that monosyllabic function words of English (like articles, conjunctions and prepositions) can undergo severe reduction, while similar monosyllabic content words (like nouns, verbs and adjectives) do not. We have begun to document this observation in a corpus of spontaneous speech (selected dialogues from the Callhome database, made up of telephone-quality speech recorded during conversations between pairs of friends or family members). Close phonetic analysis of 101 tokens of the sequence /tu/, corresponding to “to,” “two,” and “too,” in one such dialogue of more than 700 seconds, showed that the nature of the modification was (1) systematic with respect to the preceding segment, and (2) strikingly more severe than modifications in the 15 instances of content-word-initial /t/ found in the same dialogue. This work forms part of a larger project involving the investigation of the effects of prosodic structure on the phonetic modification of function words in this speech corpus.

1.3.3 Disfluencies and Speech Errors

In earlier work, we documented the constraints and patterns found in a large corpus (10,000+) of speech errors collected by writing down errors as they were heard. We have begun extending this work by developing a corpus of acoustically recorded and digitized speech errors which permits relistening at leisure, along with prosodic labelling and instrumental analysis. With Dr. Anne Cutler of the Max Planck Institute for Psycholinguistic Research at Nijmegen, the Netherlands, we have analyzed the prosodic marking of corrections of several different types of speech errors, and found that these correction patterns can be used to help differeniate word-level and sound-level errors. This is possible because, as earlier work by Cutler has shown, correction patterns for word-level errors (such as “substitution” for “advertising”) are more likely to be prosodically marked with a larger $F0$ excursion than sound-level errors (such as “shong-lort” for “long-short”).

We analyzed digitized errors that are ambiguous between word-level and sound-level because the error output is a word that differs from the target by only one sound segment, as in “moat” for “boat.” Results show that corrections of these errors behave prosodically like corrections of sound errors and not like corrections of word errors. This suggests that the difference in correction prosody does not reflect the surface character of the error (i.e., whether or not it corresponds to an existing word of English), but rather to the underlying processing level at which the error occurred (i.e., the word-processing versus the sound-processing level). This somewhat surprising result has implications for the family of cognitive models which can be proposed for speech production planning.

1.4 Models of Lexical Representation and Lexical Access

1.4.1 Overview

The development and implementation of a knowledge-based model for accessing words from continuous speech is being carried out in collaboration with Professor Carol Y. Espy-Wilson and her students at Boston University. This model has several components or modules. It contains a lexicon that is stored in memory as complexes of segments and distinctive features. Initial signal processing of the input speech stream locates a sequence of consonant and vowel landmarks corresponding roughly to the lexical segment sequence and the articulator-free features for the segments. These landmarks guide further signal processing that estimates articulator-bound features. In running speech, a talker makes modifications of the features specified in the lexicon, particularly at word boundaries. These modifications are optional and rule-governed. At a final stage in the model, a module matches the segments and features estimated from the signal to the lexical segments and features. This module must contain knowledge of the rules governing possible feature modifications. The following sections describe progress in developing some of the components of this lexical access model.

1.4.2 Automatic Location of Vowel Landmarks in Continuous Speech

In a proposed model of lexical access based on feature identification and lexical matching, an important first step is to locate the vocalic nuclei in the utterance. Locating the vowel landmarks is the initial step toward identification of the vowel features, identifying the features of adjacent consonants, and estimation of the structure of the syllables for which these vowels form the nuclei. The method that is being used to locate the vowel landmark is based in part on past work and employs measurement of low-frequency amplitude peaks, location of peaks in first-formant frequency, and first-formant amplitude. Methods of evaluating the accuracy of the vowel landmark detector are being developed, and these methods require that a clear distinction be made between underlying vowel nuclei as estimated from the lexical representation of an utterance and acoustic vowel nuclei that remain in running speech after deletion of some vowels and merging of certain vowel-vowel sequences. Evaluation of the vowel-locating algorithm with various speech databases is in progress.

1.4.3 Identification of the Voicing Feature for Obstruents

A systematic acoustic study of the cues for identification of the voicing feature for obstruents is in progress. The measures that are being examined are of three kinds: measures of the strength of glottal vibration during the constricted interval for the consonant, measures relating to the presence of glottal adduction or abduction in the vowel immediately preceding or following the consonantal interval, and evidence for increased or decreased vocal-fold stiffness in the adjacent vowel. The acoustic parameters include various measures of relative spectrum amplitude at low frequencies and of fundamental frequency in the different time intervals. Preliminary data from a database of vowel-consonant-vowel utterances show that combination of these various measures can lead to reliable identification of the voicing feature. Each of these measures separately shows some discrimination between voicing types.

1.4.4 Identification of Place of Articulation for Stop and Nasal Consonants

Theoretical models of speech production, as well as perceptual data, have led to the specification of a variety of acoustic properties that can potentially differentiate place of articulation for stop and nasal consonants. For both stops and nasals in postvocalic position, place cues are provided by the time course of the movements of formants. In initial position, the spectrum of the burst is a cue, as well as the formant movements. As a part of the development of a knowledge-based model for lexical access from continuous speech, we have been collecting data on all of these parameters from a variety of utterances ranging from spoken sentences to isolated consonant-vowel consonant utterances. Several results have emerged from this study. First, it was observed that good identification of place of articulation requires that spectral attributes be measured and averaged over carefully defined time intervals close to the consonantal closure and release. Secondly, averaging of power spectra over brief time intervals like 10-15 ms can help to produce spectra with less variability. Thirdly, by using several acoustic measures in combination, it is possible to identify place of articulation with accuracy well over 90 percent.

1.4.5 Studies of Consonant Sequences Across Word Boundaries: The Sequence /l#ð/

One of the challenges in developing a model of lexical access from running speech is the modifications that occur in some features of consonants at word boundaries. Of especial interest is the variability of the word initial consonant /ð/ when it is preceded by a word with a final coronal consonant, since this sequence occurs very frequently in English. The present study is concerned with the sequence /l#ð/ that occurs in utterances like *steal those*, or *will the*. In these sequences, it is often observed that /ð/ is produced as a sonorant consonant rather than as an obstruent, with characteristics similar to those of the preceding /l/. A goal of this kind of study is to develop acoustic analysis methods and lexical matching algorithms that permit recovery of the word sequence in spite of modifications in the acoustic realization of /ð/. A series of perceptual experiments has been carried out with word sequences of this type. When presented with word sequences such as *call those* and *call Loews*, listeners can almost always correctly identify the intended consonant /l/ or /ð/. However, if the initial word (e.g., *call*) is removed, leaving only 20-30 ms of the putative /l/ or /ð/, most of these truncated utterances are heard as beginning with /l/. Apparently the listener's interpretation of /ð/ depends upon the preceding context. Acoustic analysis of the utterances is being carried out to determine the attributes of /ð/ when it appears in this context.

1.4.6 Representation of Words in the Lexicon

There is good evidence from linguistics and psycholinguistics that speakers and listeners store words in memory as sequences of segments or phonemes. Each of these segments, in turn, is represented as a list of + or – values for a set of binary distinctive features. In the lexical access model that is being developed, a similar type of representation is being used for the lexicon. The process of accessing words involves finding a match between a pattern of segments and features derived from the speech signal to the segments and features in the lexicon.

Our initial version of the lexicon contains a relatively small number of words. As our work proceeds, we expect to expand the lexicon, and part of that expansion will include a list of affixes that can be appended to stems. This process will require that rules be written, a simple example being the plural rule for nouns in English. These rules involve changes in features or groups of features. Some of these rules can be represented in a relatively simple way if the features for the segments are represented in a hierarchical tree rather than as simple lists. This observation has provided motivation for modifying the way in which lexical items are represented. Segments in the lexicon are now represented as geometrical trees as well as by lists of features. Rules have been written for some simple morphological processes, so that the lexicon can be expanded. These rules involve manipulation of nodes (or groups of features) as well as terminal features in the geometrical trees.

Matching of signal-derived features to features in the lexicon requires that rules of a different kind be invoked—low-level rules specifying optional feature changes in segments near word boundaries in continuous speech. We expect that the use of geometrical trees for segmental representations in the lexicon will simplify the implementation of some of these rules.

1.5 Development of Facilities

We have developed hardware and software for a new PC-based digitizing system for use in projects relating to speech production. For use in the Cochlear Implant project, we have improved our setup for short-term stimulus modification experiments, in which implant users' auditory feedback is switched on and off and changes in speech parameters are recorded. Additionally, we have developed hardware that uses masking noise with normally hearing subjects, to ‘‘simulate’’ the interruption of feedback experienced by CI subjects when their implant processors are turned off. We have also been upgrading our facilities for speech analysis and synthesis, as we transfer these programs to the Linux operating system.

1.6 Publications

1.6.1 Meeting Papers

Dilley, L.C., and S. Shattuck-Hufnagel. “Ambiguity in Prominence Perception in Spoken Utterances of American English.” Paper published in the *Proceedings of the Joint Meeting of the International Congress on Acoustics and the Acoustical Society of America*, Seattle, Washington, June 20-26, 1998, vol. 2, pp.1237-38.

Dilley, L.C., and S. Shattuck-Hufnagel. “Effects of Repeated Intonation Patterns on Perceived Word-Level Organization.” Submitted to the International Congress of Phonetics Sciences, (ICPhS'99), San Francisco, California, August 1-7, 1999.

Govindarajan, K. “Latency of MEG M100 Response Indexes First Formant Frequency.” Paper published in the *Proceedings of the Joint Meeting of the International Congress on Acoustics and the Acoustical Society of America*, Seattle, Washington, June 20-26, 1999, vol. 3, pp. 2033-34.

Howitt, A.W. “Vowel landmark detection.” Abstract of paper to be presented at the Eurospeech Conference, Budapest, Hungary, September 5-9, 1999. Forthcoming.

Manuel, S.Y. “Casual Speech: A Rich Source of Intriguing Puzzles.” Paper to be presented at the International Congress of Phonetics Sciences, (ICPhS'99), San Francisco, California, August 1-7, 1999. Forthcoming.

Perkell, J., M. Matthies, and M. Zandipour. “Motor Equivalence in the Production of /ʃ/.” Paper published in the *Proceeding of the Joint Meeting of the International Congress on Acoustics and the Acoustical Society of America*, Seattle, Washington, June 20-26, 1999, vol. 4, pp. 2925-26.

Poort, K. “Stop-Consonant Production by Dysarthric-Speakers: Use of Models to Interpret Acoustic Data.” Paper published in the *Proceedings of the Joint Meeting of the International Congress on Acoustics and the Acoustical Society of America*, Seattle, Washington, June 20-26, 1999, vol. 2, pp. 1265-66.

Prahler, A.M. “Modeling and Synthesis of Lateral Consonant /l/.” Paper published in the *Proceedings of the Joint Meeting of the International Congress on Acoustics and the Acoustical Society of America*, Seattle, Washington, June 20-26, 1999, vol. 1, pp. 257-58.

Shattuck-Hufnagel, S., and A. Turk. “The Domain of Phrase-Final Lengthening in English.” Paper published in the *Proceedings of the Joint Meeting of the International Congress on Acoustics and the Acoustical Society of America*, Seattle, Washington, June 20-26, 1999, vol. 2, pp. 1235-36.

Shattuck-Hufnagel, S., and A. Cutler. “The Prosody of Speech Error Corrections Revisited.” Paper to be presented at the International Congress of Phonetics Sciences, (ICPhS’99), San Francisco, California, August 1-7,1999. Forthcoming.

Stevens, K.N. “Toward Models for Human Production and Perception of Speech." Paper published in the *Proceedings of the Joint Meeting of the International Congress on Acoustics and the Acoustical Society of America*, Seattle, Washington, June 20-26, 1999, vol. 4, pp. 2385-86.

Stevens, K.N., and S.Y. Manuel. “Revisiting Place of Articulation Measures for Stop Consonants: Implications for Models of Consonant Production." Paper to be presented at the International Congress of Phonetics Sciences, (ICPhS'99), San Francisco, California, August 1-7, 1999. Forthcoming.

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

Veilleux, N.M., and S. Shattuck-Hufnagel. “Phonetic Modification of the Syllable /tu/ in Two Spontaneous American English Dialogues." Paper Published in the *Proceedings of the International Conference on Spoken Language Processing (ICSLP'98)*, Sydney, Australia, November 1998.

1.6.2 Journal Articles

Guenther, F., C. Espy-Wilson, S. Boyce, M. Matthies, M. Zandipour, and J. Perkell. “Articulatory Tradeoffs Reduce Acoustic Variability during American English /r/ Production.” *J. Acoust. Soc. Am*. 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 and Hearing Research.* Forthcoming.

House, A.S., and K.N. Stevens. “A Longitudinal Study of Speech Production, I: General Findings.” Submitted to *J. Acoust. Soc. Am.*

House, A.S., and K.N. Stevens. “A Longitundinal Study of Speech Production, II: General Findings.” Submitted to *J. Acoust. Soc. Am.*

House, A.S., and K.N. Stevens. “A Longitudinal Study of Speech Production, III: Stop, Fricative and Nasal Consonants.” Submitted to *J. Acoust. Soc. Am.*

Kwong, K.W., and K.N. Stevens. “On the Voiced-Voiceless Distinction for Writer/Rider.” Submitted to *J. Phonetics.*

Lane, H., J. Perkell, J. Wozniak, J. Manzella, P. Guiod, M. Matthies, M. MacCollin, and J. Vick. “The Effect of Changes in Hearing Status on Speech Sound Level and Speech Breathing: A Study Conducted with Cochlear Implant Users and NF-2 Patients.” *J. Acoust. Soc. Am.* 104: 3059-69 (1998).

Massey, N., and K.N. Stevens. “Transients at Stop-Consonant Releases.” Submitted to *J. Acoust. Soc. Am.*

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.*

Perkell, J., M. Zandipour, M. Matthies, and H. Lane. “Clarity versus Economy of Effort in Speech Production: A Preliminary Study of Inter-Subject Differences and Modeling Issues.” Submitted to *J. Acoust. Soc. Am.*

Perkell, J., and M. Zandipour. “Clarity versus Economy of Effort in Speech Production: Kinematic Performance Spaces for Cyclical and Speech Movements.” Submitted to *J. Acoust. Soc. Am.*

1.6.3 Books and Chapters

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*. Merle Home, ed. Stockholm, Sweden: Kluwer. Forthcoming.

Stevens, Kenneth N. *Acoustic Phonetics*. Cambridge, Massachusetts: MIT Press, 1998.