

## XXII. SPEECH COMMUNICATION

### Academic and Research Staff

Prof. Kenneth N. Stevens	Dr. A. W. F. Huggins‡	Dr. Joseph S. Perkell
Prof. Morris Halle	Dr. Margaret Kahn	Dr. Ernest A. Scatton***
Dr. Jared Bernstein	Dr. Dennis H. Klatt	Dr. Stefanie Shattuck-Hufnagel†††
Dr. Sheila E. Blumstein*	Dr. Martha Laferriere**	Dr. Katherine Lee Williams
Dr. Margaret Bullowa	Dr. John Makhoul‡	Dr. Catherine D. Wolf
Dr. William E. Cooper	Dr. Lise Menn	Dr. Victor W. Zue
Dr. François Grosjean†	Dr. Paula Menyuk††	Barbara J. Moslin
Dr. William L. Henke	Dr. Joanne Miller†	John M. Sorensen
	Dr. Colin Painter‡‡	

### Graduate Students

Marcia A. Bush	Ursula G. Goldstein	Stephen K. Holford
Bertrand Delgutte	Howard L. Golub	George P-N. Hui
William F. Ganong III		Peter V. LaMaster

#### 1. STUDIES OF SPEECH PRODUCTION AND PERCEPTION

National Institutes of Health (Grants 5 RO1 NS04332-15 and 5 T32 NS07040-03)

Sheila E. Blumstein, Marcia A. Bush, Bertrand Delgutte, William L. Henke, Dennis H. Klatt, Colin Painter, Joseph S. Perkell, Kenneth N. Stevens, Victor W. Zue

##### a. Segmental Aspects of Speech

One of our research objectives continues to be to develop a theoretical basis for the phonetic categories or features that appear to be universal in language and to begin to delineate in detail the inventory of these phonetic features and their articulatory and perceptual correlates. During the past year, we have been concerned particularly with the features that describe place of articulation for consonants and vowels, and the features that specify tonal contrasts for vowels. Our recent work with the consonants has been acoustic analyses of a number of stop and nasal consonants in syllable-initial and syllable-final positions. Our data have provided support for earlier observations

---

\* Associate Professor, Department of Linguistics, Brown University.

† Assistant Professor, Department of Psychology, Northeastern University.

‡ Staff Member, Bolt, Beranek, and Newman, Inc.

\*\* Assistant Professor of Linguistics, Southeastern Massachusetts University.

†† Professor of Special Education, Boston University.

‡‡ Associate Professor of Communicative Disorders, Emerson College.

\*\*\* Assistant Professor, Department of Slavic Languages and Literatures, State University of New York at Albany.

††† Assistant Professor, Department of Psychology, Cornell University.

that short-time spectra sampled at consonantal release have invariant properties that distinguish fairly reliably between labial, alveolar, and velar places of articulation independent of vowel context or voicing and nasality characteristics of the consonant. These findings suggest that the auditory system is somehow predisposed to classify auditory inputs in terms of these acoustic properties. These observations have indicated that further research is needed to describe how auditory signals with these characteristics are processed in peripheral stages of the auditory system, and work along these lines is in progress, using the techniques of both psychophysics and electrophysiology.

We have been exploring various ways of describing articulatory configurations for vowels, the objective being to determine whether there are procedures for specifying articulatory states for vowels that are related in some simple, quantal fashion to the vowel features. In particular, we have been attempting, through palatography and radiography, to determine whether patterns of contact on the tongue or on the palate for the production of different vowels might provide an invariant articulatory specification of the features high, low, and back.

Research on the laryngeal features has in the past year been directed toward establishing a perceptual basis for the fundamental-frequency ( $F_0$ ) contour shapes that are observed in tone languages. Through acoustic analyses of contour shapes that are used in different languages, and through experiments that examine listener responses to contour shapes that are varied systematically in synthetic syllables, we are attempting to determine how the predispositions of the auditory system constrain the  $F_0$  contours that are used in language.

#### b. Physiology of Speech Production

Progress has been made toward establishing a capability to perform experiments in which cineradiographic and electromyographic data on speech production are obtained simultaneously. Data from these experiments will be used in conjunction with a model of the tongue for the exploration of the properties of the tongue musculature as they influence articulatory configurations and movements. Preliminary EMG records have been obtained, and methods of reducing radiographic dosage have been explored. Our computer facility is being updated to provide an interactive multichannel (physiological) data processing capability.

An experiment has been performed in collaboration with Dr. Thomas Gay (of the University of Connecticut Health Sciences Center and Haskins Laboratories) on the short-latency feedback effects of unexpected reductions in intraoral air pressure during the production of bilabial stops. In apparent conflict with a previous result, no short-latency changes in the timing of laryngeal EMG signals were found. We plan to repeat this experiment in our own laboratory.

c. Word-, Phrase-, and Sentence-Level Phenomena

The speech signal contains acoustical manifestations not only of segments and features but also of larger syntactic units. A complete theory of the speech production and perception process must account for these acoustical influences that are the result of concatenation of segments into words, phrases, and sentences. Our work in this area has included studies of speech timing in sentences, characterization of fundamental-frequency contours in sentences, and the study of certain phonological phenomena that occur when syllables and words are concatenated. This research is leading to the implementation of a model of sentence production that includes a phonological component (providing a transformation from an abstract representation in terms of segments and features, syntactic marking, and semantic information, to a detailed phonetic and prosodic representation) and a phonetic-to-acoustic transformation. Plans for improving the model include systematic evaluations of the intelligibility and naturalness of the output, and modification of the rules in areas where these evaluations show deficiencies.

2. SYNTACTIC-TO-PHONETIC CODING IN SPEECH PRODUCTION

National Institutes of Health (Grant 5 RO1 NS13028-02)

William E. Cooper, John M. Sorensen

A model of speech production has been designed to describe the translation from a speaker's syntactic representation to a phonetic output. The model is designed to account for such phenomena as (a) clause- and phrase-final lengthening, (b) optional pausing at the beginning and end of main clauses, (c) blocking of cross-word phonetic conditioning at the boundaries of major constituents, and (d) fall-rise contours of fundamental frequency at major constituent boundaries. Formal linguistic studies have been conducted to develop rules for syntactically determined pausing, and experimental studies have been conducted on each of the four phenomena mentioned above. Taken together, the experimental studies provide substantial support for the notion that a speaker's syntactic code exerts a systematic and direct influence on the phonetic properties of duration and fundamental frequency. In addition, the results of such work have permitted new inferences about the precise form of the speaker's syntactic representation. By using a controlled experimental technique which is applicable to direct hypothesis-testing, we have been able to obtain more information about the speaker's syntactic code than has previously been possible on the basis of other data.

(XXII. SPEECH COMMUNICATION)

3. STUDIES OF SPEECH PRODUCTION AND SPEECH DISCRIMINATION  
BY CHILDREN AND BY THE HEARING-IMPAIRED

National Institutes of Health (Grant 5 T32 NS07040-03)

National Science Foundation (Grant BNS76-80278)

Jared Bernstein, Marcia A. Bush, Ursula G. Goldstein, Howard L. Golub,  
William L. Henke, Margaret Kahn, Lise Menn

a. Speech and Sound Production by Infants and Children

A review of the literature has led to descriptions of the anatomical growth of the various parts of the vocal tract of children. Based on these data, a static vocal-tract model representing various levels of anatomical development is being implemented. This model will be used to explore the phonetic capabilities of children and the relationship between children's articulatory configurations for different vowels and the configurations used by adults.

In another project, we are applying modern signal-processing techniques to an analysis of infant cries, with a view toward developing improved methods of classifying the attributes of cries and individual differences in these attributes. We have obtained recordings of cries from a number of newborns, and we are collecting data on the temporal characteristics, formant frequencies, and fundamental-frequency contours for these vocalizations.

In another study, we have been gathering and analyzing speech from recordings of a number of two-to-five year olds in conversation with their parents in a structured laboratory setting. These data are being used to uncover certain principles governing the use of pitch contours in parent-child discourse, particularly the semantically-determined conditions under which the pitch range remains similar or undergoes significant changes from one clause to the next.

b. Speech Production by the Deaf

This research is concerned with identifying differences between the speech of the deaf and normal speech while attempting to isolate the causes of these differences. Speech timing and segmental patterns of a small number of deaf children have been described in detail, and one child's speech has been modeled and manipulated so that the relative contributions of different speech errors to intelligibility can be measured. An analysis of syllable and word concatenation among a larger population of deaf children during reading and spontaneous speech is being started. Another ongoing project involves transcription and acoustic analysis of segmental and suprasegmental aspects of the speech of adventitiously deaf adults. The data from this group of adults seems to indicate that, following the onset of post-lingual profound deafness, those aspects of speech

which can be monitored by rich afferent information may be less likely to deteriorate. Finally, a refined computer program has been developed for marking fundamental pitch periods. We hope that this program will be useful in the analysis of pathological phonation qualities in both children and adults.

4. ACOUSTIC STUDIES OF SPEECH SOUNDS: INVARIANT ATTRIBUTES, CONTEXT EFFECTS, AND SPEAKER DIFFERENCES

William L. Henke, Martha Laferriere, Kenneth N. Stevens, Victor W. Zue

a. Study of the Phonological Processes in English

The goal of this work is to provide, through acoustic studies, quantitative information on the variations of the properties of speech sounds in context. Whenever the variations appear to be systematic, either for all speakers or a subset of the speakers, rules are proposed to describe such phonological variations. Over the past year we have completed a study of the acoustic characteristics of medial [t, d], and we have initiated a study of the acoustic effect of vowel nasalization. In the forthcoming year, we expect to move on to other phonological processes, such as palatalization, vowel reduction, and schwa deletion.

b. Statistics of Fundamental-Frequency ( $F_0$ ) Contours in Sentences

We have been collecting long-term statistical data on  $F_0$  contours produced by a number of speakers during reading and during spontaneous speech. Three kinds of analysis of the contours are under way, representing successively more sophisticated attempts to uncover the underlying differences in the control of  $F_0$  by different speakers: (i) long-term distribution of fundamental periods; (ii) statistical properties describing an "average" breath group (a breath group being a segment of an utterance produced without pauses, and separated from adjacent segments by pauses); and (iii) statistical properties of each of several classes of breath groups.

