

# **Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech**

**John Kingston and Mary E. Beckman (editors)**  
(Cornell University and The Ohio State University)

Cambridge, England: Cambridge  
University Press, 1990, x + 506 pp.  
Hardbound, ISBN 0-521-36238-5, \$69.50;  
Paperbound, ISBN 0-521-36808-1, \$27.95

*Reviewed by*  
*Scott Meredith*  
*Apple Computer, Inc.*

*Papers in Laboratory Phonology I*, edited by John Kingston and Mary E. Beckman, contains 27 revised versions of papers and commentaries presented at the First Conference in Laboratory Phonology, held in June 1987. The editors lead off with an excellent introduction, including discussion of the motivation for the conference, explanation of its multi-disciplinary nature, and summarizations of the contributions, showing their relations both to one another and to the structure of the conference as a whole.

From the point of view of many researchers in general natural language processing, much of the background assumed and many of the questions addressed are likely to seem somewhat arcane. Nevertheless, the book is valuable on two counts: first, as an (admittedly sophisticated) overview of current practical and theoretical concerns in phonetics and phonology; and second, as an exemplary effort to integrate perspectives from radically differing "scientific subcultures" (i.e., phonetics and phonology). Traditionally, phonology has concerned itself with symbolic representations of cognitive primitives and processes that are manipulated by native speakers and hearers in conjunction with grammatical systems, while phonetics has devoted itself to instrumental analysis of the articulatory organs and the speech signal. The conference and this book are landmark efforts in attempting to reconcile these intellectual streams and lay the foundations for a unified theory of speech production and perception. As neighboring subdisciplines in all areas of NLP bring themselves into ever greater proximity by virtue of their own progress, general and methodological issues such as those addressed in this work become ever more pressing, and will require constant attention and evaluation.

## **1. Overview of the Subject Area and Contributions**

The introduction (Kingston and Beckman) explains the differing perspectives that phoneticians and phonologists bring to the presumably common goal of constructing an adequate theory of speech production and perception. The historical and contemporary tension between the empirical, physicalist perspective of phonetics, with its emphasis on instrumental analysis, and the cognitive orientation of phonology, with its appeals to speaker/hearer intuition and links to broader linguistic theory, is clearly, even entertainingly, described. The conference from which the book emerged was motivated by an attempt to reconcile these intellectual streams, or at least begin a dialog. The book mirrors the structure of the conference, consisting of theme papers with several commentaries on each (in most cases) by researchers of differing orientation.

Three general areas are addressed:

**Suprasegmental Phenomena (Pitch and Duration).** Questions raised here include the representation of tone and intonation (What are the cognitive primitives of pitch representation? How are these entities reflected in the speech signal? To what degree are they uniquely recoverable by hearers? What relations exist between cognitive and physical analytical categories for languages of different surface typologies?). Papers in this subgroup include: Inkelas and Leben on Hausa intonation; Ladd on metrical representation of pitch register; commentary by Clements; Silverman and Pierrehumbert on certain kinds of high tone accents in English; commentary by Bruce; Kohler on the relative importance of segmentally conditioned, local pitch perturbations with respect to global contours and percepts; commentary by Silverman.

Duration effects are also addressed (comparison of domains of application of F0 effects with those of duration effects; 'head-based' versus 'edge-based' effects; relation of syntactic constituency to patterns of segmental duration variation; speaker- versus hearer-based models; psychological reality). Papers in this subgroup include: Beckman and Edwards on prosodic constituency; Selkirk on the relation of syntactic structure to prosodic effects; commentary by Fowler raising methodological and phonetic issues; and Cutler commenting on the speaker's versus the hearer's perspective.

**Relation Between Phonological and Phonetic Representations.** Questions addressed here are quite similar in general character to those raised in the preceding section on suprasegmentals, i.e., the exact specification of phonological categories and their combinatory possibilities, the relevant parameters for phonetic measurement and evaluation, and how exactly the latter can be predicted from the former (for a given utterance), or the former justified on the basis of the latter (for theory in general). Papers in this area include: Hertz's description of the Delta speech synthesis control formalism and implementation; Ohala on phonetic motivations for common phonological patterns; Pierrehumbert on the prospects for a clear separation between levels of phonetic and cognitive representation; and commentary by Ohala.

**Descriptive Systems for Speech Segments, and Their Relation to Production and Perception.** Questions addressed here include the phonetic and phonological basis of typical patterns of segmental and syllable structure; levels of abstraction in phonetic accounts; representation of articulatory gestures and their coordinated timing and control; and, as above, the possible link between representation systems that have been found useful for addressing phonological concerns (such as typological or historical patterns) and those that have been found relevant in addressing phonetic questions of speech production and perception. Papers in this area include: Clements on segmental features that determine segment positioning within a syllable; Browman and Goldstein on an articulatory model of phonology and phonetics; commentary by Steriade; a radical phonetic perspective from Ladefoged; Kingston on a principle of articulatory control of segments; commentary by Ohala; commentary by Goldstein; Keating on a model of coarticulation; commentary by Stevens; commentary by Fowler.

## 2. Discussion

This is not an easy book. Even professional phoneticians and phonologists will have to work at getting the most from it. That this should be the case is, in a sense, unsurprising, given that one of the primary motivations for the conference was bridging the perceived chasm of mutual indifference, ideology, and methodology between the two professional subcultures. Therefore, phoneticians may find some of the theoretical assumptions, notational and categorical systems, etc., to be irrelevant or apparently arbitrary, while some phonologists may be slowed down by the statistical presentations,

waveform displays, etc. Readers with a more casual interest may be stymied by ubiquitous tacit assumptions concerning historical and methodological issues that have informed both fields since their inception. The arguments presented are for the most part tightly reasoned and very well structured, but assume a high degree of interest and motivation on the reader's part. The conference organizers clearly knew what they were doing when they made up the invitation list: here are the leaders of the two fields, all in one place.

It is important for all researchers and engineers working on any aspect of speech, and perhaps natural language processing in general, to be aware of the existence of this book, for two reasons.

First, on the practical side, there are questions raised and solutions proposed that will be increasingly relevant as speech engineering advances beyond current rudimentary speech synthesis and recognition systems into areas where near-human performance is expected. In order for engineering in these areas to progress, research and development may have to be structured according to, or at least informed by, the questions raised in this book. This is particularly true in the area of suprasegmentals or prosody. Recent research aimed at providing a basis for improved synthesis and recognition systems (e.g., Wightman et al. 1992) has begun to link concrete data analysis efforts with conceptual structure derived from phonological and phonetic theory of exactly the sort presented in depth in the first section of this book. Thus understanding (or at least awareness) of the issues presented here will both increase appreciation for such emerging efforts and suggest new, principled directions for research and development.

Second, on the conceptual side, this conference gives a unique foretaste of an approach to an increasingly interesting problem: attempting to reconcile radically different perspectives on essentially integrated phenomena. While the situation in phonetics and phonology may be extreme, both in terms of their expected total affinity and actual near-total lack of practical working relations, roughly comparable situations exist in other pockets of natural language research as well, e.g., symbolic versus connectionist accounts of grammatical phenomena, syntactic and semantic theory versus NLP systems engineering, and introspective versus psycholinguistic evidence for aspects of syntactic and semantic theory. Many other examples could be adduced. This book is exemplary in suggesting an approach to productive dialog and a holistic perspective.

Finally, while the book's content is difficult, the editors and publishers have gone to some trouble to provide an attractive presentation: the book is nicely organized, printed in a consistent typeface, with numerous clear charts and tables, and with separate indices for names and subjects.

#### Reference

Wightman, Colin W.; Shattuck-Hufnagel, Stefanie; Ostendorf, Mari; and Price, Patti J. (1992). "Segmental durations in the

vicinity of prosodic phrase boundaries." *Journal of the Acoustical Society of America*, **91**, 1707-1717.

*Scott Meredith* is a Research Scientist at Apple Computer, Inc., Advanced Technologies Group, Speech and Language Technologies. He did his doctoral work at MIT in prosodic phonology, and his current research interests include text-to-speech prosody, multilingual text-to-speech systems, and related issues in speech recognition. Meredith's address is: Apple Computer, Inc., 20525 Mariani Avenue MS 76-4E, Cupertino, CA 95014. email: meredith.s@applelink.apple.com