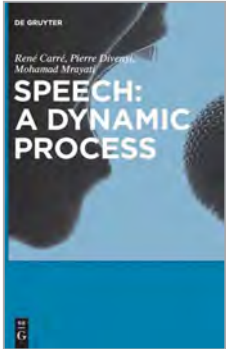


Book Review

These reviews of books and other forms of information express the opinions of the individual reviewers and are not necessarily endorsed by the Editorial Board of this Journal. – Philip I. Marston, Book Review Editor

Speech: A Dynamic Process



Authors: René Carré, Pierre Divenyi, Mohamad Mrayati

Publisher: De Gruyter, Berlin, Germany, 2017, 214 pp.

Price: £79.56 (Hardcover); £63.96 (Kindle)

ISBN: 978-1-501-510601

The world's many tongues vary greatly in their sound patterns, lexicons, and grammar. Despite such variation, all languages share certain properties in common. One of especial interest is the intricate interleaving of consonants and vowels in the articulatory stream, instantiated as a low-frequency modulation in the speech waveform. Such fluctuations are the result of articulatory movement of the lips, tongue and jaw, and which serve as the acoustic foundation of a sophisticated information-bearing system.

Why speech is so configured is rarely addressed in the scientific literature.

Speech: A Dynamic Process examines the question from several vantage points through a clever combination of modeling, theory, and empirical studies.

The book begins with a consideration of two approaches to scientific inquiry — induction (with its focus on data collection, statistics, modeling and prediction) and deduction (emphasizing “universal” principles, “logic,” formal modeling, and verification). These complementary perspectives are used to integrate the trifecta of speech production, acoustics, and perception within a unified theoretical framework.

This philosophical introduction is followed by an historical review of speech research, with a focus on the latter half of the 20th century that sets the scene for the vocal-tract modeling work discussed in later chapters.

The models range from the primitive — a simple tube — to the more elaborate, incorporating far greater degrees of freedom.

A key issue considered is how a vocal tract model can be deformed in such a way that results in maximal acoustic variation with minimal vocal effort.

The authors propose a “Distinctive Region Model,” (DRM) in which “privileged” acoustic trajectories (i.e., formant patterns) serve as “coding primitives.” In their view, information is encoded in articulatory trajectories rather than as a sequence of quasi-static “targets.” “Locus equations,” which compute the slope and duration of formant transitions, are shown to reliably distinguish among vowels. Because these trajectories are shaped by the interaction of vocalic and consonantal context, the framework is essentially one in which the syllable forms a fundamental unit of speech production and perception.

Although the model is restricted to vowels and stop consonants, it appears to capture much of the speech signal's dynamic properties. One of the most intriguing properties of the DRM is its ability to produce a broad range of speech sounds with a relatively small number of control parameters, an insight of potential interest to speech clinicians and engineers.

A later chapter focuses on perceptual aspects, including empirical studies of processing dynamic speech sounds. This work is then linked back to the modeling and production studies discussed in the earlier chapters. A case is made for the human auditory system's specialization for speech, especially for processing communication signals in the presence of background noise and other forms of acoustic interference.

The final chapters are more philosophical in tone, presenting the case for a dynamic theory of speech based on melding the inductive and deductive approaches discussed earlier. The theory is also examined from the perspectives of linguistics, psychology, engineering, and evolutionary theory.

Speech: A Dynamic Process is a thought-provoking book that examines the speech signal in many interesting ways. However, the theoretical framework proposed has its

limitations. The models discussed are, by necessity, simplifications of a very complex system that is still not well understood. The modeled data are drawn mostly from highly simplified speech, whose acoustic properties differ in many respects from the patterns observed in naturally spoken language. Another limitation is the framework's difficulty in explaining how extremely distorted speech remains intelligible under challenging listening conditions in which much of the waveform has been excised or heavily masked. How does the brain manage to track formant trajectories under such conditions? Also unaddressed are the roles of dynamic visual-speech cues and memory in decoding the speech signal.

Despite such limitations, the book is likely to interest researchers wishing to examine speech from an unusual vantage point that's both original and thought-provoking. The bibliography is excellent and comprehensive. The text is clearly written, and the figures informative and insightful.

Review by:

Steven Greenberg

Silicon Speech

17270 Greenridge Road

Hidden Valley Lake, CA 95467

(Published online December 10, 2018, The Journal of the Acoustical Society of America, 144(6).)



"In Tune"

by Alex Tolstoy

Watercolor
©2019 by Alex Tolstoy,
all rights reserved.

We are delighted that ASA member Dr. Alexandra Tolstoy will provide some of her original water colors (all with acoustics themes) for *Acoustics Today* from time to time. You can see more of Alex's paintings at atolstoyart.com.