ACOUSTICS2008/2369
Mathematical Evidence For Motor Theories of Speech Perception

Gordon Ramsay
Haskins Laboratories, 300 George Street, New Haven, CT 06511, USA

Theories of speech perception have often proposed, controversially, that recovery of phonological information from the acoustic signal necessarily requires implicit knowledge of the physical processes underlying speech production. Most of the evidence supporting this hypothesis has been derived empirically from experiments in speech perception. In this paper, we show that many of the basic ideas underlying motor theories of speech perception can also be derived mathematically from first principles, using classic results in stochastic nonlinear filtering theory. Two key results, the Duncan-Mortenson-Zakai and Fujisaki-Kallianpur-Kunita theorems, show that the optimal state estimator for any partially-observed nonlinear stochastic dynamical system always takes the form of a "matched filter", which is itself a nonlinear stochastic dynamical system, the structure of which mimics, and resonates with, the structure of the original. Interpreting this in the context of speech perception, recovery of phonological information from sound produced by a human vocal tract necessarily involves construction of an internal model of the processes implicated in speech production. A key prediction, which we explore, is that any such model need not reproduce all the details of these processes, but is only required to predict the lawful conditional correlation between gesture and sound.