Journal of Phonetics 64 (2017) 34–50

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Special Issue: Mechanisms of regulation in speech, eds. Mücke, Hermes & Cho

Exertive modulation of speech and articulatory phasing

Sam Tilsen

Dept of Linguistics, 203 Morrill Hall, Cornell University, Ithaca, NY 14853, United States

ARTICLE INFO

Article history:
Received 1 April 2016
Received in revised form 17 February 2017
Accepted 1 March 2017
Available online 21 March 2017

Keywords:
Motor control
Articulation
Articulatory timing
Coordination
Coupled oscillators
Attention

ABSTRACT

An articulatory study was conducted to investigate whether fluctuations in exertive mechanisms (attention, effort, motivation, arousal, etc.) have a global effect on articulatory control systems. Participants in the experiment produced an articulatory pattern 400 times, attempting to do so as consistently as possible. Evidence for global exertive modulation was obtained in the form of widespread correlations between variables associated with biomechanically independent systems such as phonation, linguo-labial coordination, and head movement/posture. Analyses of movement timing autocorrelation showed evidence for random walk-like dynamics on short timescales and equilibrium dynamics on long timescales, along with evidence for low- and high-exertion states of production. An extension of the coupled oscillators model of articulatory coordination is presented to account for these phenomena.

1. Introduction

Speech involves a variety of systems which interact on a range of spatial and temporal scales. Most of the relevant systems—i.e. ensembles of neurons, individual speakers, social networks—are typically far from equilibrium. In other words, non-equilibrium states are the norm. Nonetheless, the interactions between non-equilibrium systems may produce emergent patterns which exhibit equilibrium-like behaviors. Whenever we observe hints of such patterns, it is important to study the relevant systems in detail, because they greatly constrain our models and provide insight into underlying control mechanisms.

This study focuses on one such pattern, the symmetric temporal displacement observed in the initiation of consonantal and vocalic articulatory movements. Unconditioned, spontaneous fluctuations in this timing pattern were collected in an experimental paradigm in which participants produced a single target response 400 times, attempting to do so as consistently as possible.

Two hypotheses were tested: an *exertive modulation hypothesis*, which holds that exertive mechanisms (e.g. attention, effort, arousal) induce random walk-like dynamics in a variety of independent speech motor control systems, and an *equilibration hypothesis*, which holds that equilibration mechanisms constrain articulatory control systems on long timescales, in effect confining the exertive random walk in a potential. The exertive modulation hypothesis was supported by positive lag-1 autocorrelations of response variables on short timescales, along with pervasive correlations between outputs of biomechanically independent motor systems. The equilibration hypothesis was supported for some participants by decreases in autocorrelation on longer timescales. It was also observed that variation in a model-derived proxy for exertive force was associated with differing profiles of variance and covariance in articulatory timing, suggestive of a contrast between high- and low-exertion regimes. These findings are important because they provide a new basis for understanding the mechanisms involved in speech production.

1.1. Equilibrium vs. random walk behavior of systems

The concept of equilibrium arises in many domains. For example a mechanical equilibrium refers to a situation in which the net force on an object is zero, a chemical equilibrium describes an equivalence of forward and reverse chemical reaction rates, and a population equilibrium describes a biological system with stable predator and prey populations. In all of these cases, if the equilibrium is stable, there is a cost (in energy or in some other quantity) for deviations from the equilibrium, with greater deviations being more costly. When

E-mail address: tilsen@cornell.edu

http://dx.doi.org/10.1016/j.wocn.2017.03.001
0095-4470/© 2017 Elsevier Ltd. All rights reserved.

fluctuations from the environment perturb a system from its equilibrium state, the system subsequently returns to the equilibrium.

In contrast, a system with random walk behavior does not have an equilibrium or steady state. Generically, in a random walk the state of the system changes at each time step with a random displacement from the previous state. In the absence of any external forces, there is no cost for changing states, and hence if we wait long enough, we can expect the system to be arbitrarily far from its starting point. Clearly the systems which are responsible for controlling speech production cannot be governed solely by random walk dynamics, but it does not follow that there are no random walk-like components to their behavior. One possible scenario is that speech control systems have equilibrium states, but fluctuations in the nervous system add a random walk-like component to system dynamics. Evidence for this scenario can only be found if the responses of systems to departures from equilibria are relatively slow. Hence to investigate these possibilities, we must examine speech patterns over an extended period of time.

A useful statistical approach to investigating these ideas is to estimate the autocorrelation functions of system outputs. A Gaussian white noise process exhibits zero autocorrelation at all lags, because any state of the system is independent from all previous states. This is a generic property of a system which quickly returns to its equilibrium when perturbed—as long as the return to equilibrium is faster than the measurement period, successive observations will be uncorrelated. In contrast, a random walk tends toward a lag-1 autocorrelation of one, because each state depends on the previous one. No equilibrium is present in a random walk. (See Box, Jenkins, Reinsel, & Ljung, 2015; Chatfield, 2016; Sethna, 2006 for introductions to time series analysis and random processes.)

Many systems of interest may have more complicated structure, such as a random walk in an external field or a random walk with an external noise source. To assess whether observations might be generated by such a system, it is useful to conduct analyses over a range of timescales by applying a coarse-graining procedure to the observation sequence. The coarse-graining used here involves averaging observations over non-overlapping windows of time, thus integrating out short-timescale fluctuations. The lag-1 autocorrelation *scaling function* shows how the lag-1 autocorrelation changes as a function of the size of the averaging window. Fig. 1 (right) shows mean and ±1 standard deviation for autocorrelation scaling functions of several different types of processes. Over all coarse-grain timescales (τ) the Gaussian noise process tends toward a lag-1 autocorrelation of zero, while the random walk tends toward a lag-1 autocorrelation of one (see Appendix A.1 for details).

Unlike Gaussian noise and a random walk, more complicated systems exhibit lag-1 autocorrelations that vary substantially with analysis timescale. For example, a random walk with an external Gaussian noise (e.g. measurement noise) converges to random walk-like autocorrelation when coarse-grained, but is less than one on short timescales. Alternatively, a random walk in a quadratic potential has a random walk-like autocorrelation on short timescales but eventually converges to Gaussian noise-like equilibrium dynamics on longer scales. Combining a random walk, external Gaussian noise, and a quadratic potential results in an autocorrelation scaling function which increases at short timescales, peaks at some intermediate scale, and decreases on longer timescales. We will see this same profile in the autocorrelation scaling functions of articulatory timing measures, which suggests that models of articulatory control require mechanisms for both random walk- and equilibrium-like dynamics.

1.2. Symmetric displacement in articulatory timing and the coupled oscillators model

In many languages a pattern of articulatory timing called the *C–center effect* is observed (Browman & Goldstein, 1988; Hermes, Mücke, & Grice, 2013; Marin & Pouplier, 2010; Tilsen et al., 2012), which involves the symmetric displacement of the initiations of consonantal movements from the initiation of a vocalic movement in syllables with complex onsets. As schematized in Fig. 2 (left), the initiations of C_1 and C_2 constriction gestures in a C_1C_2V syllable are equally displaced in

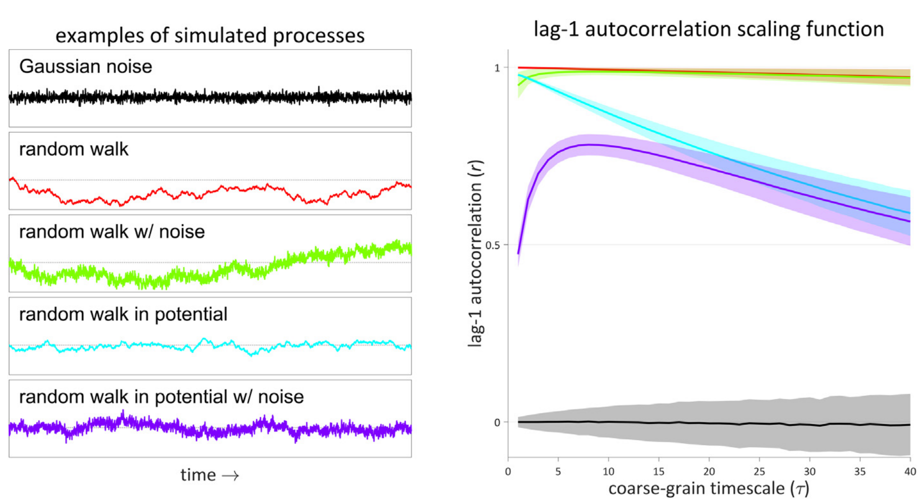

Fig. 1. Comparison of autocorrelation scaling functions of several random processes. (Left) example time series. (Right) Lag-1 autocorrelation scaling functions. Filled areas are ±1 s.d. of autocorrelation functions from 1000 simulations of 5000-sample observation sequences.