Journal of Phonetics 64 (2017) 8–20

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

The relation between the continuous and the discrete: A note on the first principles of speech dynamics

Khalil Iskarous

Department of Linguistics, 3601 Watt Way, 301 Grace Ford Salvatori, University of Southern California, Los Angeles, CA 90089-1693, United States

ARTICLE INFO

Article history:
Received 27 May 2016
Received in revised form 8 April 2017
Accepted 9 May 2017
Available online 12 June 2017

Keywords:
Speech dynamics
Dynamical systems analysis
Principle of least squares
Gestures

ABSTRACT

The goal of this paper is to show how dynamical theories of phonetics and phonology bridge the dualistic gap between discrete phonological descriptions and continuous phonetic descriptions. By delving into the first principles of dynamics, it is shown that dynamical theories do not assume separate sets of principles to describe discrete and continuous aspects of a system. Rather, the discrete description is shown to predict the continuous one, using the concept of a differential equation, which is thoroughly explained. Linear and nonlinear differential equations are introduced using a discrete approximation, and then used to show how phonological contrast has been accounted for using dynamical systems analysis. A dynamical recurrent neural network model of word formation is then discussed to show how linguistic plans for words are serialized and coordinated into motoric word plans for different articulatory systems in the vocal tract. Furthermore, it is shown that many aspects of the discrete, time-invariant phonological description can be predicted from observed variable continuous phonetic functions, using the principle of least squares and recurrent neural networks.

1. Introduction

When we consider the entity /waj/ (*why*), in some dialects of American English, there is a sense in which it contains one thing, three things, many things, or be considered continuous. This is because /waj/ is a single linguistic entity, a word, with a particular meaning in these dialects, as well as a combination of three segments, chosen from a few such segments in the language that can be serialized and overlapped in different ways to compose meaningful entities, and it can also be considered as a combination of many minimally contrastive units. But any actual production [waj] of this entity is continuously variable. Of course nothing is special about this word or these dialects. Any description of a spoken or signed natural language is confronted with the problem of how to combine the continuous and discrete aspects of language. The traditional approach to this problem is a dualistic Cartesian description. The discrete aspects are considered to be part of a cognitive description of the sound aspects of a language, phonology, while the continuous aspects are considered part of a physical description of the continuous acoustic, articulatory or visual aspects of the language, phonetics. The two disciplines interface, but they are fundamentally different in what they describe and in their theoretical structure, with the former centered around logic and grammar and the latter centered around physics and biology. In the last several decades, however, several researchers have challenged this dualism through a *dynamical* description of language (Browman & Goldstein, 1985; Browman & Goldstein, 1989; Byrd & Saltzman, 1998; Byrd & Saltzman, 2003; Elman, 1995; Fowler, Rubin, Remez, & Turvey, 1980; Gafos & Benus, 2006; Iskarous, 2016; Jordan, 1986; Kelso, Tuller, Vatikiotis-Bateson, & Fowler, 1984; Kelso, Vatikiotis-Bateson, Saltzman, & Kay, 1985; Lindblom, 1983; Munhall, Ostry, & Parush, 1985; Perrier, Ostry, & Laboissiere, 1996; Roon & Gafos, 2016; Saltzman & Munhall, 1989; Saltzman, Nam, Krivokapić, & Goldstein, 2008; Smolensky, Goldrick, & Mathis, 2014; Sorenson & Gafos, 2016; Tilsen, 2016; Tuller, Case, Ding, & Kelso, 1994). This theoretical approach has received empirical support through a wide variety of studies (e.g., Browman, 1994; Browman & Goldstein, 1988; Byrd, Tobin, Bresch, & Narayanan, 2009; Chen, Chang, & Iskarous, 2015; Goldstein, Pouplier, Chen, Saltzman, & Byrd, 2007; Iskarous et al., 2013; Katsika, Krivokapić, Mooshammer, Tiede, & Goldstein, 2014; Krivokapić, 2014; Marin & Pouplier, 2010; Pouplier, Marin, & Kochetov, 2015; Pouplier & van Lieshout,

E-mail address: kiskaro@gmail.com

http://dx.doi.org/10.1016/j.wocn.2017.05.003

2016; Shaw & Gafos, 2015; Vatikiotis-Bateson & Kelso, 1993). Even though these papers offer different viewpoints and hypotheses about the structure of spoken language, they all emerge from a common dynamical framework, Dynamical Systems Theory, initially developed by Isaac Newton to describe many phenomena of the physical world. This framework combines the description of the discrete and the continuous into one entity, the differential equation. And that is the essence of why this framework is also applicable to simultaneously describing phonology and phonetics.

It is felt by the author that at this juncture of the development of the dynamical approach to phonetics and phonology, it is worthwhile to revisit its first principles, and to re-argue its applicability to the resolution of the antagonism of the discrete and the continuous. It is also felt that further empirical development of the theory will rely more and more on a deep understanding of the basic mathematical structure of dynamical systems theory, therefore, a novel tutorial introduction to the differential equations of dynamics will be provided, using a discretization of the theory, reducing it to arithmetic calculation. This approach makes the ideas stand out, and does not rely on the symbolic manipulations of the calculus, which are not necessary for grasping the workings of the theory. This approach is based on the belief that the differential equation core of dynamical systems analysis is not an *implementational* issue, which would be safe to ignore unless one is simulating the theory. Rather, the form and the content of these equations is the heart of the theory, and is at the basis of its adequacy for linguistic description, extension to account for additional phenomena, and empirical falsifiability.

This will be done in several steps. First it will be shown how linear dynamical systems theory allows for the prediction of a global continuous trajectory in time from a discrete description, and how this development underlies the Browman and Goldstein (1985), Browman and Goldstein (1989) and Saltzman and Munhall (1989) theory of the continuous motoric expression of discrete contrasts in language. Second, it will be shown how nonlinear dynamical systems theory conceives of the binary phonological contrast (Tuller et al., 1994), such as voice, and the phonological intention (Gafos & Benus, 2006) to select one of the featural specifications, such as [voiceless]. Third, it will be shown how Jordan (1986)'s theory of serial order allows for relating one discrete cognitive entity into a serially ordered as well as overlapped set of motoric time series, and how this theory was incorporated by Saltzman and Munhall (1989) and Rubin et al. (1996) as a theory of intergestural coordination, which is related to the oscillator theory of intergestural coordination (Browman & Goldstein, 2000; Goldstein et al., 2007; Nam, Goldstein, & Saltzman, 2009). Fourth, it will be shown how the principle of least squares and recurrent neural networks can be used to infer the discrete description from the continuous one. This possible inversion of dynamical descriptions will be argued to be essential, if these theories are to be regarded as intrinsically related to and inferable from observed data, rather than just theoretical abstractions.

2. Linear dynamical systems and contrast

One of Isaac Newton's greatest discoveries was the idea that the fundamental laws of nature are differential equations (Arnold, 1986). He considered it such a great discovery that he made an anagram of it, to claim priority, and sent it to Leibniz. Differential equations are descriptions of the relationship between a dependent variable's value, also called the state of the system, and the value of that variable at some infinitesimally close value(s) of the independent variable. The wide range of applicability of dynamical systems stems from the plethora of concrete and abstract variables that can be interpreted as states of dynamical systems, e.g. position of a particle, temperature at a point in space and time, valuation of a stock, activation of a neuron, lip aperture (LA), or formant value. One of the simplest such differential equations describes the state x of a system in the following way: $\frac{dx}{dt} = -kx$, where k is some value $0 \leqslant k < 1$. This law can be read: the difference between the state x of a system at some point in time t, and the state $x(t + dt)$ of that same point in the infinitesimally close future, normalized by the change in time, is equal to some known invariant property of the system – k times the value of the current state $x(t)$. The law is approximately true, with the extent of the approximation determinable quantitatively, if the increments of time dt and state dx are small but not infinitesimal, changing them to Δt and Δx as finite small numbers. The differential equation is said to have been discretized. The approximate law can then be written as $\frac{\Delta x}{\Delta t} = -kx$, where Δ indicates a difference. We can therefore think of the relation as $\frac{x_{future} - x_{present}}{\Delta t} = -kx_{present}$, which can be used for calculation as:

$$x_{future} = -k\Delta t\, x_{present} + x_{present} = (1 - k\Delta t)x_{present} \tag{1}$$

If some initial state is known, that can be taken as the current $x_{present}$ on the right hand side of (1). And x_{future} can be calculated, if we know k and we have chosen Δt. That previous future state could then be taken as the current state, from which a new future state can be calculated, etc. For instance, if k were $\frac{1}{3}$, and the state x at the initial time were 100, and we were to choose Δt as 1, then the position after one time point would be 66.667 (since $(-\frac{1}{3} * 1 * 100) + 100 = 66.667$). If we now take this new state value to be the present one, then the future value would be 44.44, etc. Since the law is (approximately) true at all time points, the value of the position will approach 0 at infinite time, which is easily checkable by iterating (1) for this example. These values of x for this example can be seen as circles in the topmost curve of Fig. 1a. This function approximates a global *continuous* (since Δt can be made infinitesimally small) function solution to the differential equation, based on the *discrete* value of k and the differential relation.[1] The other curves in the figure show the evolution of x for different initial values (80, 60, etc.). We say that 0 is a stable equilibrium value, since all nearby curves head towards it. The word *equilibrium* refers to the fact that if x were actually 0, $\frac{dx}{dt} = -kx = 0$, and there is no longer any change. 0 acts as a *goal* for the system, achieved by the system, regardless of where it starts. Fig. 1b shows curves for the same initial values,

[1] Finite or discrete approximation to differential equations is a well-understood branch of numerical analysis (e.g. Strang, 2007). Moreover, over the last 100 years the calculus has seen generalizations of the notions of the calculus to discrete and discontinuous domains, e.g., Lebesgue integration and nonsmooth analysis, fields based on which routine discretization can be justified theoretically. Further details on differential equations and their discretization can be found in Boyce and DiPrima (2012).