Journal of Phonetics (1990) **18**, 465–477

On attributing grammars to dynamical systems

J. B. Pierrehumbert

Department of Linguistics, Northwestern University, 2016 Sheridan Road, Evanston, IL 60201, U.S.A.

R. T. Pierrehumbert

Department of the Geophysical Sciences, University of Chicago, 5734 S. Ellis Ave. Chicago, IL 60637, U.S.A.

1. Introduction

Who could dispute the obvious fact that language as physically implemented is describable by a continuous dynamical system? Even digital computers, viewed as physical devices, are continuous dynamical systems. The physics of speech production involves continuously deformable viscoelastic elements coupled to continuously fluctuating fluid flow. As physiological experiments have attempted to track language into the brain, it has turned out that all available measurements exhibit continuous variation. This is not surprising, since individual neurons are themselves complex dynamical systems.

The remarkable thing is that the output of this complicated dynamical system is characterized by a considerable degree of regularity at the discrete level. Discrete, or "qualitative", regularities in language sound structure are the primary objects of study in phonology. Phonology concerns not only the inventory of elements which languages manipulate, but also the grammar which specifies how they may be combined. As the field has progressed, it has showed a constant give-and-take between claims about the inventory and claims about the grammar. Adjusting the theory of the inventory can lead to a better characterization of the grammar, and adjusting the grammar can lead to a better characterization of the inventory. For example, studies of the phonological behavior of geminates and homorganic clusters have led both to a revision in the theory of the units (with autosegments displacing phonemes), and a revision in the principles of combination (Hayes, 1986; McCarthy, 1986; Ito, 1988). Regardless of one's opinion of the utility of discrete representations, they are an indisputable observed property of the system; although many mathematically possible dynamical systems do not exhibit grammatical structure, evolution has not adopted such a system for human communication. Therefore, any valid phonological theory based on dynamical systems needs to be able to reproduce these regularities, just as Newtonian dynamics is able to reproduce Kepler's laws of planetary motion.

Among schools of thought on this subject, at one extreme is the position that phonological theory should deal directly with the dynamics, the surface regularity emerging more or less automatically as a macroscopic property of the system much as classical bulk thermodynamics emerges from the statistical mechanics of interacting particles. At the other extreme is the position that the continuous

0095–4470/90/030465 + 13 $03.00/0

dynamics is so hopelessly far from the deep regularities of language, that it is pointless to do anything other than deal with the discrete representation directly. Our own position is somewhere between the two. We believe there is a great deal of productive work to be done at the interface between continuous dynamics and discrete phonology. Work in phonology promises to provide insight into the dynamics, and work in dynamics promises to illuminate the phonetic basis of phonology, which is broadly apparent but which has proved exasperatingly difficult to characterize in detail. Given the amount of progress in identifying the phonetic basis for the segmental inventory (e.g. Stevens, 1989; Wood, 1985), we would particularly emphasize the need for work on the phonetic basis of the grammar. However, we also feel that a good deal more precision must be brought to the discussion of how continuous and discrete representations of the behavior can be reconciled. Far from being a peripheral issue, the rigorous relation of grammatical and dynamical descriptions is one of the most central and challenging of current research issues.

It is certainly useful to be reminded from time to time that non-linear systems have complex behavior (as in Browman & Goldstein's position paper and commentary in this issue, 1990*a,b*), but this is where the hard work begins, not where it ends. In this commentary, we shall see a bit of where the mathematics leads us when we attempt to go further, by discussing established means of formally associating grammars with dynamical systems. These established methods involve positing two disparate levels of representations, one qualitative and one quantitative, and then identifying a deep relationship between them. The quantitative representation does not supplant the qualitative one, which in some cases provides answers to questions which are intractable from the quantitative perspective. We also examine where established methods fall short of the requirements for phonological applications, and show where some of the difficulties lie in effecting the necessary extensions. In the end, we hope to have provided an idea of what a rigorous dynamically-based theory of phonological regularity might look like, even though at present most of the key ingredients are missing.

2. Categories and dynamics

There are two parts to the problem of forging a link between a continuous dynamical system and the discrete representation familiar to linguists. One part is the derivation of the discrete categories from the continuous signal. These categories form the inventory of symbols in our formal language. The other part is to derive *from the dynamics* a set of rules governing the allowable ("well-formed") sequences of symbols. This is the grammar of the language. These two parts are related since it is possible for a continuous system to suggest several different categorical interpretations, with some serviceable for a rigorous treatment of the grammar and others not. Therefore work on phonological categories, such as the quantal theory of speech, needs to be embedded in a grammatical treatment.

Let us consider a simple example, illustrating what we mean. The first order system

$$\frac{dx}{dt} = -(x - x_0(t)) \tag{1}$$

yields solutions $x(t)$ which tend exponentially toward the variable target $x_0(t)$. In this system one is getting neither the categories nor the grammar from the dynamical system. All this is incorporated in the externally specified sequence of target positions $x_0(t)$. The general solution to (1) is

$$x(t) = x(0)\mathrm{e}^{-t} + \int_0^t x_0(\tau)\mathrm{e}^{\tau-t}\,\mathrm{d}\tau \tag{1a}$$

Models like (1) are thus equivalent to a linear filter of the input signal $x_0(t)$. The same comment applies to the articulatory model employed in Browman & Goldstein (see the two papers in this issue 1990*a*,*b*), which leads to a somewhat more complicated filter. We are not saying that such representations are without merit; indeed they have a history in the treatment of F_0 (e.g. Öhman, 1967; Fujisaki *et al.*, 1979), and Browman & Goldstein have made a substantial contribution by drawing out their implications for the more complex domain of segmental allophony. However, they do not fully bridge the gap between dynamics and phonology. "Task dynamics" as it is most typically carried out to date takes *discrete* imputs and produces *continuously variable* outputs (see e.g. Figure 6 in Saltzman & Munhall, 1989). To quote Saltzman & Munhall:

> Currently these activation patterns are not derived from an underlying implicit dynamics. Rather, these patterns are specified explicitly "by hand", or are derived according to a rule-based synthesis program...(p. 343).

In Browman & Goldstein (1990*b*), the grammatical information supplied to the dynamics specifically includes not only the segmental sequence, but also the prosody. This is the opposite of what is required. Rather than accounting for phonology on the basis of dynamics, it amounts to a procedure for phonetic implementation of the phonology. Such studies can make an important contribution to phonology when they show that the phonetics accounts for variations previously thought of in terms of low-level phonological rules. The phonology can be simplified when it sheds the burden of describing regularities which are difficult to treat grammatically (see Pierrehumbert, 1980, and Pierrehumbert & Beckman, 1988, as well as Browman & Goldstein's discussion of reduction). However, the approach does not do away with phonology because it says nothing about what sequences of elements are well-formed. Jordan's recent work on "serial dynamics" described in Saltzman & Munhall (1989) is in a somewhat different category, as it has the potential for autonomous generation of regularities in the control sequence. In Section 2 we will have a bit more to say about where this line of work fits into the general scheme of things.

In contrast to (1), consider the damped non-linear oscillator equation

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = a(t)x(1-x^2) - (1+\mathrm{e}^x)\frac{\mathrm{d}x}{\mathrm{d}t} \tag{2}$$

where a is a generalized spring constant.[1] This equation has three fixed points $x = 0, \pm 1$. Suppose for the moment that a is time-independant. Then, when $a < 0$, $x = 0$ is a stable fixed point ("point attractor") and $x = \pm 1$ are unstable fixed points ("point repellors"). When $a > 0$, the situation is reversed and $x = \pm 1$ are the point

[1] For $a < 0$, the undamped system is prone to finite-time blowup when x becomes larger than 1. The non-linear damping coefficient helps to suppress this distraction.

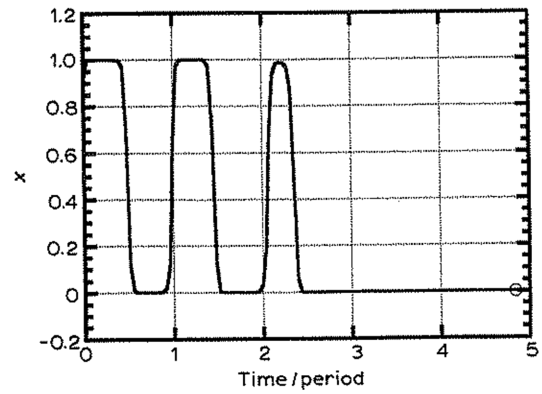


Figure 1. Time sequence for Equation (2) subject to periodic variation of the control parameter.

attractors of the system. In this case, the categories come out of the dynamics, in the form of the inventory of fixed points. If the system starts close to $x=1$ with $a>0$ and a is gradually reduced to a negative value, then x will make a nearly discontinuous transition to $x=0$. However, the grammar governing the possible sequences of transitions between fixed points still is put in externally, through the control parameter $a(t)$. Mind, we are not saying that category formation in phonology will turn out this simply. In fact, models of this sort, involving an exchange of stability of fixed points, run into serious problems as soon as one attempts to discuss the transition between states. The difficulty arises because the system approaches its point attractor exponentially closely during the time the control parameter a retains a fixed sign; hence it tends to get "stuck" in the corresponding position, and the longer a remains, say positive, the longer it must be held negative in order to cause a transition to the $x=0$ state. This is an undesirable feature from the linguistic standpoint, as is demonstrated by the following example. Suppose that a varies in time with the form

$$a(t) = \gamma \cos(\omega t) \tag{3}$$

One would like a periodic variation of the control parameter to lead to an alternation of states (as in an extended CV alternation such as /bababa.../). An integration of (2) with $\gamma = 0.5$ and $\omega = 0.05$ is shown in Fig. 1. The system does indeed show an alternation of states for the first few periods, but then it falls into the $x=0$ state and stays there indefinitely. This kind of behavior is typical, in the sense that one need not choose the parameters very carefully in order to see it. From this example we learn that the mere presence of bifurcation in a system does not assure a natural path to the analysis of the dynamics of transitions between states.

3. Markov partitions and symbolic dynamics

We will now turn to the analysis of systems for which both the categories and the grammar proceed from the dynamics in a natural way. The *symbolic dynamics* formalism provides the clearest instance in the dynamical systems literature of the success of such a program. An introduction to symbolic dynamics may be found in Guckenheimer & Holmes (1983).

The logistic map

$$x_{n+1} = rx_n(1 - x_n) \tag{4}$$

provides a good illustration of the application of symbolic dynamics. Analysis of chaos in this system is often associated with the work of May, who certainly helped to bring the phenomenon to the attention of biologists and investigators in diverse other fields. However, quantitative analysis of the chaos and its physical implications dates back at least to Lorenz (1964), who in turn was building on the work of Ulam (1960). In our example, we concentrate on the case $r = 4$, for which (4) is a map from the interval $0 \leq x \leq 1$ to itself, and the *orbits* (i.e. sequences x_0, x_1,...) are chaotic.

We divide the interval into three subintervals

$$\begin{aligned} a&: \; 0 \leq x \leq 0.1464466094 \\ b&: \; 0.1464466094 \leq x \leq 0.5 \\ c&: \; 0.5 \leq x \leq 1 \end{aligned} \tag{5}$$

assigning the symbol a to times when x_n is in the first interval, and so forth. Note that the "objects" which are set up by this partition of the phase space are different from the kind of "objects" considered in the previous example; in addition, they cannot be viewed as a macroscopic or average property of the system as Browman & Goldstein propose. When (5) is applied, the system (4) generates a string of symbols. Can the set of possible output strings be characterized by a grammar dealing only with the discrete representation? Because of the careful way the partition (5) has been constructed, the answer is affirmative.

From (4) it follows immediately that the following transitions between adjacent symbols are possible: $a \rightarrow \{a,b\}, b \rightarrow c,\; c \rightarrow \{a,b,c\}$. This defines a *finite state grammar* for a language composed of sequences of the three symbols. Trivially, any output of (4) is a well-formed sequence in this language. It is a more remarkable fact that, with the partition (5), there is a *one-to-one correspondence* between symbol sequences of the language and orbits of (4); for any well formed symbol sequence, there is a *unique* initial condition x leading to that sequence (modulo multiplicity associated with uniform shifts in the sequence). Hence, the finite state grammar completely characterizes the set of strings that can be output by (4). This convenient state of affairs will not hold for just any partition we might pick. Partitions with the desired property are known as *Markov partitions*.

The inner workings of the Markov partitions are shown in Fig. 2 for the periodic sequence *cabccabc*... . Any initial value of x in the range roughly between 0.96 and 1.0 leads to the sequence *ca*. A somewhat narrower subinterval leads to *cabc*, and a narrower one still leads to *cabcc*. Only initial conditions in a very narrow interval surrounding $x = 0.98$ leads to *cabccabc*. The interval collapses to a point as we increase the number of periods over which we require the output to track the target sequence. This behavior, which is critical to making the rigorous association between the grammar and the dynamical system, is at the same time somewhat problematic from a linguistic point of view. It says that the entire future course of the signal is encapsulated in the initial physical condition. It is as if a regular (though of course non-linear) relationship were proposed between the details of the first

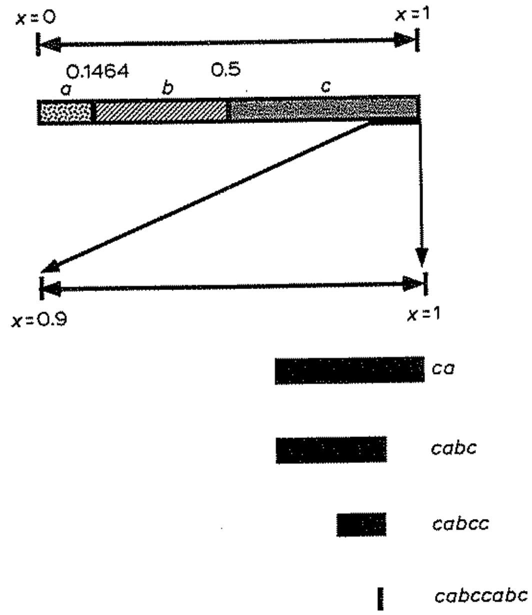


Figure 2. Markov partitions for the logistic map.

vowel production and the entire sequence of upcoming segments. Even if the infinite sequences of symbolic dynamics were truncated or conflated because of error propagation, the appropriateness of this approach is unclear, because much recent work on phonology and phonetics has emphasized the locality of access to phonological structure (cf. Ito, 1988; Pierrehumbert & Beckman, 1988).

The grammatical structure associated with (4) shows that the output of chaotic systems is not equivalent to random noise. There appears to be some confusion surrounding this point in the task dynamics literature. For example, Saltzman & Munhall (1989), in discussing strange attractors,[2] say: "For nonrepetitive and nonrandom speech sequences, such attractors appear clearly inadequate" (p. 356). In fact, chaos is special precisely because it is structured but non-periodic. Among dynamical systems, or even within a given system for various parameter values, the degree of order can vary continuously in a manner that has been made precise and quantitative by Crutchfield & Young (1989). Thus, Saltzman & Munhall's attempt to characterize the output of "connectionist models" as fundamentally different from strange attractors is not particularly meaningful. The models in question may form a particular class of non-linear dynamical system, with perhaps a particular kind of chaotic output. They are nonetheless amenable to the same kinds of analysis applied to dynamical systems in general, and present the same difficulties. Whether

[2] It is worth noting at this point that most of the phenomena generally referred to by this term have not been proved to be either "strange" of even "attractors" in the strict mathematical sense. In two decades of applications of chaos in the physical sciences, there do not appear to have been any major problems caused by glossing over this gap, and so we feel reasonably safe in continuing to use the term loosely. In any event, it is well established that invariant sets that are not attractors can have a profound influence on the long term time evolution of the system. The periodic orbits we refer to in our symbolic dynamics example are all *unstable*. However, they are dense in the phase space, and so the strange attractor is in a sense made up of bits and pieces of these orbits of varying lengths.

their special form makes the analysis of structure in the output easier or harder remains to be seen. The going is apt to get difficult, as the neural nets under consideration involve rather large numbers of interacting non-linear elements. Models of this complexity are proposed with good reason, because the evidence plainly indicates that phonology is physically implemented as wetware in the brain. This leads us to the question of how well symbolic dynamics carries over to systems with large numbers of degrees of freedom.

For maps on the line, such as the logistic map, Markov partitions can be constructed in a systematic way. Geometrically, this is possible because the boundaries between partitions are points, which one can slide back and forth on the line like beads on a wire until the appropriate mathematical conditions are met. For systems with more than one degree of freedom, the situation is far less satisfactory. Consider the map

$$x_{n+1} = 20x_n y_n e^{-x_n y_n} \tag{6a}$$

$$y_{n+1} = g(y_n + ax_n) \tag{6b}$$

where $g(z) = z$ for $z < 2$ and $g(z) = 1$ for $z > 2$. This is a two degree-of-freedom system in the sense that two variables (the initial x and y) suffice to determine the future course of the system. For fixed x, (6b) would lead to a sawtooth oscillation of y with a period that increases as ax decreases. For fixed y, (6a) would lead to a set of behaviors similar to the logistic map. For small y, $x = 0$ is a stable fixed point; as y is increased, sequences of stable periodic orbits become possible and ultimately chaos sets in. In the coupled system, the control parameter of (6a) is set internally by the position in the oscillation of y, and the average level of excitation of x in turn affects the period of the oscillation. In a linguistic context, one might think of (6b) as an internal clock generating gestural control sequences, with the control values translated into output by articulators modeled (metaphorically!) by (6a).

The attractor for this system for $a = 0.05$ is shown in Fig. 3. It is constructed by taking an initial condition, iterating (6) 11 000 times, and plotting the last 10 000 points only, so as to give the system time to settle down into its long-term behavior. The attractor is clearly a geometrically complex entity, made up of a large (actually uncountable) number of contorted, closely-spaced sheets. Its curious geometry is why it is called a "strange attractor", though the term is a misnomer, because such

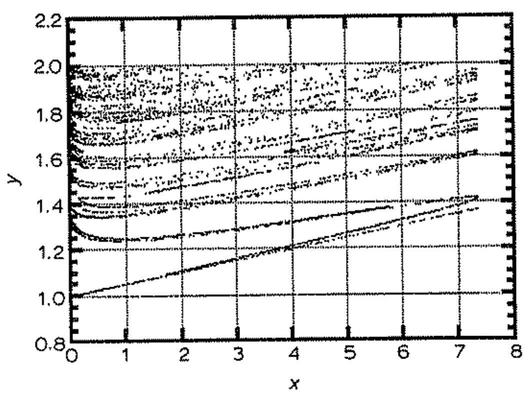


Figure 3. Attractor for Equation (6), based on 10 000 iterations.

things are in fact quite typical. In order to obtain a Markov partition, we would have to cover the strange attractor by a tiling of the plane which satisfies certain delicate mathematical conditions concerning which subregions of each region map into each other region. Rather than varying positions of a small number of points, as for one-dimensional maps, we must vary the set of *curves* bounding the regions. We are not saying this is impossible, but it is clearly of a completely different level of difficulty from the one-dimensional case. Only a handful of two-dimensional systems have been successfully analyzed in terms of symbolic dynamics (e.g. Smale, 1963; Meiss & Ott, 1986; Rom-Kedar, 1990), and these by rather idiosyncratic methods that do not readily generalize. Leaving aside the question of construction, there are not even any generally applicable results on the class of two-dimensional systems for which Markov partitions *exist*. They are known to exist for a special class of n-dimensional systems satisfying a technical requirement known as *hyperbolicity*, but hyperbolicity is difficult to prove, and in any event many systems of physical interest are known to be non-hyperbolic, including the logistic map in the parameter range discussed above.

This does not mean that the output of such systems is random. We have constructed a "linguistic corpus" by iterating (6) 100 000 times and assigning the symbol a to $x \leq 1.35$, and b to $x > 1.35$. The two symbols appear with roughly equal frequency in the corpus, but the corpus exhibits striking regularities and preferences for certain patterns. For example, among all possible sequences of 12 symbols, the string *babaabaaabab* is the most common, appearing 589 times. However there are *no* occurrences of strings with *aaaa* as a substring. The more repetitive sequence *baabaabaabaa* occurs only 53 times, even though it has none of the unfavored *aaa* substrings. The sequence *baababaabab*, obtained by transposing the first two "feet" of the most common string, does not occur at all.

The news from dimension-2 is rather dismal, but it gets even worse: without even considering the high-dimensionality implied by the astronomical numbers of degrees of freedom of the neurons and mechanical elements involve in producing speech, simple constructs familiar from control theory suffice to render the system infinite-dimensional. Consider the delay differential equation

$$\frac{dx}{dt} = -x_{(t-T)} \tag{7a}$$

where T is a fixed time lag and the notation on the right-hand side denotes that x is to be evaluated at time $t-T$ rather than at time t. This is an *infinite-dimensional* system, because the behavior of $x(t)$ over the entire interval $-T \leq t \leq 0$ must be specified in order to determine the future course of the system. Splitting up time into discrete chunks of length T, (7a) can be regarded as a map on the space of all integrable real functions on the interval $0 \leq t \leq T$. Function spaces are very high-dimensional creatures indeed. Of course, (7) is a linear system, and its continuous family of solutions can be obtained systematically by employing the usual arsenal of methods applicable to linear systems. The ice gets thinner for non-linear systems, as the reader can verify by experimenting with the "delayed logistic equation"

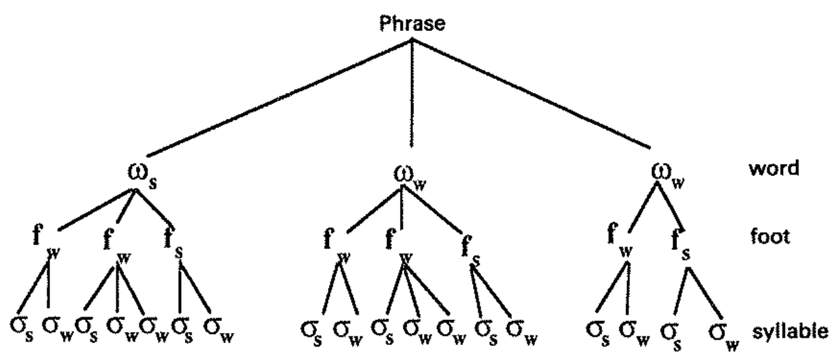
$$\frac{dx}{dt} = rx_t(1 - x_{(t-T)}) \tag{7b}$$

It remains to be seen whether some kind of extension of symbolic dynamics together with a suitable underlying dynamical system can account for the observed phonological regularities. Regardless of the outcome, from symbolic dynamics we come away with a renewed appreciation of the utility of multiple levels of representation. Even when the underlying continuous system is known exactly (as in the logistic map), introducing a discrete level of representation enables one to derive properties of the system that are difficult to obtain directly on the basis of the continuous dynamics. For example, the symbolic dynamics of the logistic map immediately reveals that the system permits infinitely many (unstable) periodic orbits, which can have arbitrary length. Counting the number of periodic orbits of a given length can give a characterization of such important properties of the continuous signal as its entropy. If discrete representations have proved essential for simple systems where the underlying dynamics is known, they are doubly so for systems like phonology where the discrete regularity has been established by observation, but the underlying dynamics is unknown and of uncertain complexity.

4. Beyond finite-state grammars

Markov partitions provide a tidy answer to the problem of rigorously associating a grammar with a dynamical system, but finding them is a challenging problem. At the same time, it is not clear that they are the right tool, as they have the undesirable limitation of yielding only finite-state grammars. Given the conspicuous importance of hierarchical structure in language, one would like to replace them with something more powerful. Although there has been some work on the relation of more general grammars to dynamics (e.g. Crutchfield & Young, 1990), the work is in its infancy, compared to symbolic dynamics.

Metrical phonology is the branch of phonology dealing with hierarchical structure. It posits trees, which characterize grouping and relative prominence, and serve to coordinate the autosegmental tiers. A somewhat oversimplified prosodic tree is shown in Fig. 4. In this example, the phrase has been expanded as three phonological words, the words are expanded into feet, and the feet into syllables. Within each group at each level, one element is singled out as strong. In this example, we have adopted the position of Selkirk (1980, 1984), according to which each level in the tree corresponds to a prosodic type with a constellation of



determinate phonological properties. This position is also adopted in the treatment of surface representation in Pierrehumbert & Beckman (1988). An important corollary is that the phonological grammar has no recursion, since it contains no phrase structure rules with the same symbol on the right and on the left. Apparent recursion in the lexical phonology is described by principles relating the representations for different items to each other, (cf. Halle & Vergnaud, 1988). Thus the phonological grammar could in principle be expanded into a finite state grammar. However in practice this would result in an explosion of the state space and a loss of explanatory power.

The scientific inadequacy of a finite state treatment of phonology can be appreciated by considering the kinds of phenomena which led to the development of metrical phonology. Some of these phenomena pose specific problems for an effort to recast phonological grammar in terms of dynamic systems theory.

Well-formedness constraints which are enforced within groups but not across group boundaries comprise one major class of evidence for hierarchical structures. Such constraints are found at all levels. For example, sequences of consonants which are impossible within a syllable may be possible over a syllable boundary. A major class of rules within Lexical Phonology applies only within words, not across word boundaries. One consequence of the English stress rules is that the beginning of a word can display at most one reduced syllable before the first full syllable is reached; disregarding word boundaries, a much greater number of reduced syllables can precede a full syllable. The intonational grammar rules out the occurrence of pitch accents after the main stress of a phrase; but pitch accents can follow in a subsequent phrase. A general consequence is that if two phonological structures are well-formed, their concatenation is well-formed, modulo any constraints imposed at higher levels.

This generic property of phonology is not easily met for strings generated by partitions of dynamical systems. In examples such as (5), which are rigorously describable by finite state grammars, grammatical constraints are defined only over adjacent terminal symbols and there is no way to define boundaries which prevent the constraints from being enforced. In the mathematical experiment (6), on the other hand, the strings *aaa* and *aa* individually occur frequently in the corpus, but their concatenation does not occur at all. Here we have a candidate for a higher-level constraint, but there is no formal support for a general formulation of it. To get round this difficulty, one might propose a clock hierarchy. In the course of its vacillation each clock would generate signals to reset the dynamical system for the next lower level and destroy information prior to the boundary. Clock hierarchies have indeed been proposed by researchers in the area of motor control (e.g. Shaffer, 1981; Vorberg & Hambuch, 1978). However, the empirical arena for such proposals has been the details of timing as complex actions are carried out, rather than specification of what action sequences are well-formed. Dynamical modeling of the boundary occurrences, the means by which the reset process is implemented, and the precise nature of the information destroyed at word boundaries, constitute challenging and interesting topics for research.

A related problem is the transmission of properties up and down the hierarchy. Inheritance from above is exemplified in Pierrehumbert & Beckman (1988), who show how pitch range as defined at two levels of phrasing in Japanese is inherited in the production of individual tones. Inheritance of stress from above constrains which syllables can be assigned pitch accents. The dependence of higher levels on lower levels is illustrated by the very common occurrence of quantity sensitive stress rules. In languages with a quantity sensitive stress rule, the rule assigning feet to syllable sequences starts a new foot when it comes to a heavy syllable (a syllable with more than one element in the rhyme), even if the adjacent foot is not of maximal length. The delicacy of this problem may be illustrated by English, in which a quantity sensitive stress rule determines the location of the last stress of the word and a quantity insensitive rule assigns earlier stresses in the word (Hayes, 1980, 1982). In general, the upwards propagation of information is highly restricted, taking the form of constraints relating well-formedness of a higher level to structural properties of a lower level. If arbitrary information could be passed up to a higher node and then broadcast from there to all subordinate nodes, the restrictive force of the hierarchical representation would clearly be lost.

Downward inheritance can be incorporated in dynamical models in a straightforward way. One first assigns a variable or group of variables to each level of the hierarchy (y for foot-level, x for syllable-level, and so forth). Then, to implement inheritance one stipulates that the equations governing each level be independent of variables lower down in the hierarchy, but dependent on variables higher up. In our foot/syllable example, the system would be of the general form $y_{n+1} = G(y_n)$, $x_{n+1} = H(y_n, x_n)$. Even if they can be made to yield the correct grammar, such prescriptions suffer from having the hierarchy hard-wired into the system. They cannot constrain the set of possible hierarchical relationships, nor shed light on how hierarchies are acquired by the language learner. Adding some upward inheritance to this picture is problematic. If it is done in a straightforward way, by allowing the foot-level dynamical system to be $y_{n+1} = G(x_n, y_n)$, one is back to a general dynamical system like (6), and the hierarchical structure is lost. Identifying the class of dependencies that correspond to the *limited* kind of upward propagation described in the preceding paragraph is a challenging problem. It suggests the necessity of some combination of look-ahead schemes, short-term buffering, and local parallelism.

A third line of evidence for the hierarchical structures of metrical phonology is based on the abstraction of time that they afford. This abstraction is used to provide a limited capability for well-formedness to depend on future events. Specifically, well-formedness can depend on the categorical description of those future events which are rendered accessible by the hierarchical representation. An obvious example is the rhythm rule of English, which adjusts the stress location (or possibly the pitch accent location, see Beckman, DeJong & Edwards, 1988) in anticipation of an upcoming stronger stress. It is common for word-level stress rules to assign metrical feet sequentially from the end of the word to the beginning, in the sense that the location of the first strong syllable exhibits a systematic dependence on the characteristics of the end of the word. The parametric theory of stress rules treats the direction of assignment of metrical feet as a language specific parameter; the extent to which time is abstracted in the theory is reflected in the fact that the rest of the descriptive apparatus is the same regardless of the direction of assignment of feet (Halle & Vergnaud, 1987). Now, it is broadly apparent that metrical phonology is founded on the characteristics of rhythmic motor activity, just as distinctive feature theory is founded on the dimensions of contrast made available by vocal tract acoustics. However, the way that time is abstracted in metrical rules shows that the metrical description is not merely parasitic on the production dynamics. Instead, metrical phonology points to a cognitive capability to construct descriptions of dynamics and use them in planning operations; the need for several levels of representation appears inescapable.

5. Concluding remarks

We have argued in favor of multiple levels of representation in the theory of language sound structure, embracing continuous dynamics and discrete regularity. First, discrete phonological regularity, such as the existence of grammars, is a clear observed property of the phenomenon. Any new approach based on dynamics must be reconciled with regularity at this level. The connection between the continuous and discrete (or macroscopically observable) behavior in the linguistic case is more subtle than the analogous relation in statistical thermodynamics, as the discrete behavior does not jump out as a macroscopic average of the continuous system (contrary to the suggestion of Browman & Goldstein, 1990*b*). Second, we have shown that grammars can indeed be associated with dynamical systems, but that there is a big gap between the current development of the mathematics and what needs to be accomplished for linguistic applications. Required extensions include methods for treating higher dimensional systems and hierarchical structure. Finally, for non-linear systems, knowing the underlying dynamical equation does not provide the key to its behavior; it just provides a convenient means of simulation, which may take the place of experiment. In many cases where the dynamics is known exactly, mathematicians have been forced to invent a discrete level of representation in order to figure out what is going on. In "discrete linguistics" we skip the dynamical middleman, and go directly to trying to understand the behavior in terms of discrete representations obtained from observing the phenomenon itself (language produced by speakers). If any of these phenomena are ever to be accounted for on another level by a continuous dynamic representation, knowledge of the discrete dynamics will no doubt prove an essential clue to the reconstruction.

References

Beckman, M., DeJong, K & Edwards, J. (1988) Intonational and rhythmic correlates of stress clash, *Journal of the Acoustical Society of America* **83**, S1, S113.
Browman, C. & Goldstein, L. (1990*a*) Gestural specification using dynamically-defined articulatory structures, *Journal of Phonetics*, **18**, 299–320.
Browman, C. & Goldstein, L. (1990*b*) Representation and reality: physical systems and phonological structure. *Journal of Phonetics* **18**, 411–424.
Crutchfield, J. P. & Young, K. (1989) Inferring statistical complexity, *Physical Review Letters*, **63**, 105–108.
Crutchfield, J. P. & Young, K. (1990) Computation at the onset of chaos. In *Complexity, entropy, and physics of information*, (W. Zurek, editor). Addison-Wesley.
Fujisi, H., Hirose, H. & Ohta, K. (1979) Acoustic features of the fundamental frequency contours of declarative sentences in Japanese, *Annual Bulletin*, **13**, 163–173. Research Institute of Logopedics and Phoniatrics, University of Tokyo.
Guckenheimer, J. & Holmes, P. (1983) *Nonlinear oscillations, dynamical systems and bifurcations of vector fields*. New York: Springer Verlag.
Halle, M. & Vergnaud, J.-R. (1988) *An essay on stress*. Cambridge, MA: MIT Press.
Hayes, B. (1980) A metrical theory of stress rules. Ph.D dissertation, MIT. Distributed by Indiana University Linguistics Club, Bloomington.
Hayes, B. (1982) Extametricality and English stress, *Linguistic Inquiry*, **13**, 227–276.

Hayes, B. (1986) Inalterability in CV phonology, *Language*, **62**, 321–351.
Ito, J. (1988) Syllable theory in prosodic phonology. New York: Garland.
Kiparsky, P. (1985) Some consequences of lexical phonology, *Phonology Yearbook* **2**, 85–138.
Lorentz, E. N. (1964) The problem of deducing the climate from the governing equations, *Tellus*, **16**, 1–11.
McCarthy, J. (1986) OCP Effects: gemination and antigemination, *Linguistic Inquiry*, **17**, 207–263.
Meiss, J. D. & Ott, E. (1986) Markov tree model for transport in area preserving maps, *Physica*, **20D**, 387–397.
Öhman, S. (1967) Word and sentence intonation: a quantitative model. STL-QPSR 2-3, Royal Institute of Technology, Stockholm.
Pierrehumbert, J. & Beckman, M. (1988) *Japanese tone structure* Linguistic Inquiry Monograph Series 15, Cambridge, MA: MIT Press.
Rom-Kedar, V. (1990) Transport rates of a class of two-dimensional maps and flows. *Physica D*, **43**, 229–268.
Saltzman, E. L. & Munhall, K. G. (1989) A dynamical approach to gestural patterning in speech production, *Ecological Psychology*, **1**, 333–382.
Selkirk, E. (1980) The role of prosodic categories in English word stress, *Linguistic Inquiry*, **11**, 563–603.
Selkirk, E. (1984) *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
Shaffer, L. H. (1981) Performances of Chopin, Bach, and Bartok: studies in motor programming, *Cognitive Psychology*, **13**, 327–376.
Smale, S. (1963) Diffeomorphisms with many periodic points. In *Differential and combinatorial topology* (S. S. Cairns, editor), pp. 63–80. Princeton: Princeton University Press.
Stevens, K. N. (1989) On the quantal nature of speech, *Journal of Phonetics*, **17**, 3–45.
Ulam, S. M. (1960) *A collection of mathematical problems*. New York: Interscience publishers.
Vorberg, D. & Hambuch, R. (1978) On temporal control of rhythmic performance. In *Attention and Performance VII* (J. Requin, editor), Hillsdale, NJ: Lawrence Erlbaum.
Wood, S. (1985) The acoustical significance of tongue, lip, and larynx maneuvers in rounded palatal vowels, *Journal of the Acoustical Society of America*, **80**, 391–401.