Dynamics of Phonological Cognition

Adamantios I. Gafos

Department of Linguistics, New York University & Haskins Laboratories

Stefan Benus

Department of Computer Science, Columbia University

June 7, 2005

Abstract

A fundamental problem in spoken language is the duality between the continuous aspects of phonetic performance and the discrete aspects of phonological competence. We study two instances of this problem from the phenomenon of voicing neutralization and vowel harmony. In each case, we present a model where the experimentally observed continuous distinctions are linked to the discreteness of phonological form using the mathematics of nonlinear dynamics.

Keywords: grammar, subsymbolic computation, phonology, phonetics, nonlinear dynamics

Dynamics of Phonological Cognition

1 Introduction

A primary aim in the cognitive science of language is to discover the computational principles underlying our ability to speak. A major problem is how to relate the symbolic or discrete aspects of our speaking competence to their continuous manifestation in terms of vocal tract action. Traditionally, the study of these two aspects of speech has been pursued under separate domains, the symbolic aspects being the domain of *phonology* and the continuous aspects being the domain of *phonetics*.

How is the discreteness of phonological systems related to the continuity of their phonetic substance? This question defines the problem of the phonology-phonetics *relation*, the central theme of the 'laboratory phonology' research community (Beckman & Kingston 1990:1). It is also an instance of a broader question in cognitive science, namely, the question of how to relate the low-dimensional, discrete aspects of cognition to the high-dimensional, detailed aspects of performance. For parallel research on vision see Haken (1990), on coordination in biological action see Turvey (1990), and on agent-environment interactions see Beer (1995).

At the heart of this problem one meets a major methodological challenge. Computation is embedded in a continuously varying environment. To understand computation we must use inferences based on surface, performance data extracted from specific contexts. Abstracting away from contextual or environmental factors requires an understanding of how computation adapts to different contexts which in turn assumes an understanding of computation. Kosslyn (1978) has aptly dubbed this "the inference problem" for cognitive psychology and spells out in pedagogically useful detail an approach to dealing with this problem. The key idea in Kosslyn's method is that both competence and the nature of the mapping from competence to surface data need to be explicitly studied.

Following Kosslyn’s approach, we think that progress on our theme question is best achieved when the relation between discrete and continuous aspects of phonology-phonetics is formalized explicitly for representative cases. In this paper, we propose models of the relation between continuity and discreteness for two language-particular but nevertheless generalizable phenomena. The underlying formal basis of our modeling is the mathematical theory of nonlinear dynamics.

Our approach departs from the majority view on the relation between phonology and phonetics in two important respects. According to that view, the relation between discrete and continuous aspects of phonology-phonetics is best conceived in terms of a translation between two different formal languages, the language of discrete symbols and the language of continuous parameters. This is the view in the background of most current work on language and cognitive science in general (see the notion of transducer in Fodor & Pylyshyn 1981 and Harnad 1990). The first key difference between our view and the translational view is that in our approach there is no translation. Nonlinear dynamics enables us to integrate the discreteness and continuity of the phonology-phonetics system within the same formal description. The second difference is a corollary of the first. In the dominant view of how phonology and phonetics are put together, phonology is derivationally antecedent to phonetics. The discrete aspects of our speaking competence are the domain of a computational system manipulating symbolic representations. The output of this computation is then translated to continuous physical properties of an articulatory and acoustic nature. Our approach, instead, does away with the temporal metaphor of precedence between the qualitative and the quantitative, without losing sight of the essential distinction between the two.

The paper is organized as follows. Section 2 covers background notions from dynamics. Section 3 discusses applications of dynamics to phonological representations. Sections 4, 5 discuss applications of dynamics to organizational principles of linguistic grammars. Section 4 addresses the interaction of environmental variables with the phonological grammar in the phenomenon of voicing neutralization. The model proposed therein maintains the discrete aspects of grammar but also accounts for the continuous phonetic variation due to changes in environmental parameters. Section 5 addresses the problem of how low-level spatial phonetic properties of vowels relate to the high-level phonological behavior of suffix choice in the phenomenon of vowel harmony. We conclude in section 6 with a summary of the main points.

2 Background notions from dynamics

A fundamental property of any phonological system is that it maintains a macroscopically stable form under varying environmental conditions. Yet, the system is also flexible in that category boundaries can be shown to vary smoothly and categories can be changed by continuous scaling in stimulus (environmental) parameters. For changes in perceptual categories with scaling of stimulus parameters see Liberman *et al*. (1957), Eimas *et al*. (1971), and Repp & Liberman (1987), and for changes in production as speech rate is increased see Stetson (1951), Kelso *et al*. (1986), and de Jong *et al*. (2001). If, as we claim, the units of phonological structure and their organizational principles are dynamical in nature, then stability and change follow as natural consequences of their nonlinear dynamics. Stability, one consequence of nonlinearity, refers to the capacity of dynamical systems for maintaining preferred modes in the face of variations in environmental parameters or perturbations due to noise. Change, the other consequence of nonlinearity, results when environmental variables are scaled beyond certain critical values.

To express these theoretical ideas precisely and model experimental results, we employ concepts and tools from a branch of mathematics known as nonlinear dynamics. To review the basic notions, we begin by a general formulation for the simplest class of dynamical systems, namely, deterministic, first-order, autonomous dynamical systems (Percival & Richards 1982). These systems are described by a differential equation, $\dot{x} = f(x)$ where x is the state of the system and $f(x)$ is the force, a nonlinear function of x. For first-order systems, the force can be expressed as a function of the derivative of a potential $V(x)$, $\dot{x} = f(x) = -dV(x)/dx$. An intuitive grasp of the dynamics of our state variable x can be inferred by examining geometric properties of $f(x)$ or its related potential $V(x)$. For example, consider the behavior of a particle placed in the potential of Fig. 1, and assume that the position of this particle, its x value, corresponds to the state of the system. The points x_k where $f(x_k) = 0$ represent states of equilibrium – if our particle is placed initially at x_k it remains there for all time. Such points are called *fixed points*. There are two types of fixed points, *stable* and *unstable*. Stable fixed points correspond to the minima of the potential $V(x)$ – see x_1, x_3 in Fig. 1. Around these points the force function $f(x)$ is a decreasing function of x, or intuitively, the arrows on the x axis in Fig. 1, which show the *flow*, point towards that point. Unstable fixed points correspond to the maxima of the potential $V(x)$ – see x_2 in Fig. 1. Around such points, $f(x)$ is an increasing function of x, and the arrows of the flow point away from that point. Stable fixed points are also called *attractors*, and unstable fixed points *repellers*.

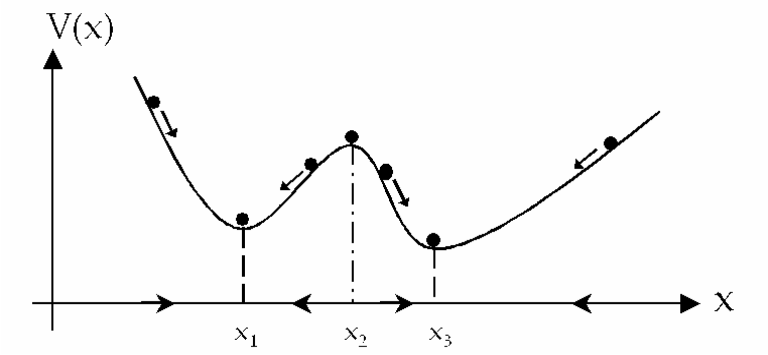


Fig. 1 Potential $V(x)$, attractors x_1, x_3, repeller x_2 of a simple dynamical system.

Fig. 1 represents the assumption that x draws values from two recognizably distinct parts of its state space (the state space is the entire x axis). It thus describes qualitatively distinct states of the system indexed with x, or in other words, it describes a dimension of macroscopic order. For this reason, it is called an *order parameter* (Haken 1977). In behavioral patterns, order parameters have the quality of *dynamic stability* (term borrowed from B. Goodwin 1970). This means that attractive states exhibit small fluctuations around their mean values (x_1, x_3 above). Fluctuations are due to noise. Noise is present due to the organizational complexity of behavior, that is, the fact that behavior involves parallel activity of distinct faculties at different hierarchical levels. At a high level, any behavior can be described by a few parameters whose dynamics are coupled to lower level subsystems controlling the more specific components. For phonological cognition, this level corresponds to the macroscopic phonological parameters (e.g. place and degree of oral constriction, larynx state, velum position). The lower levels correspond to the neuronal, aerodynamic and myodynamic subsystems controlling vocal tract action. This coupling between the two levels necessarily introduces noise in the dynamics of the high level parameters. Following Haken (1977), we describe noise as a small, random perturbation force pushing the high level parameter x back and forth randomly. Mathematically, noise enters the dynamics of high level parameters as in (1), with the additional factor representing Gaussian white-noise of strength Q. This is the general form of a nondeterministic or stochastic dynamical system.

$$(1) \qquad \dot{x} = f(x) + Noise = -dV(x)/dx + Q\sqrt{\xi_t}$$

The presence of noise introduces indeterminacy in the behavior of the order parameter x. Consequently, we can only compute the probability of finding x within a given region of values. This probability is described by the probability density function *p(x)* multiplied by the length of the region. For any nondeterministic, first-order dynamical system, there exist analytical methods allowing us to compute the probability density function by finding a stationary solution to the Fokker-Planck equation (Haken 1977, Freidlin & Wentzell 1984). An example of a probability density function corresponding to a bistable potential is shown in Fig. 2. It can be seen that the probability of finding the system around the mean states of the two attractors is quite high. As we move away from the mean states, the probability of finding the system at some other region decreases quickly but it may not be zero.

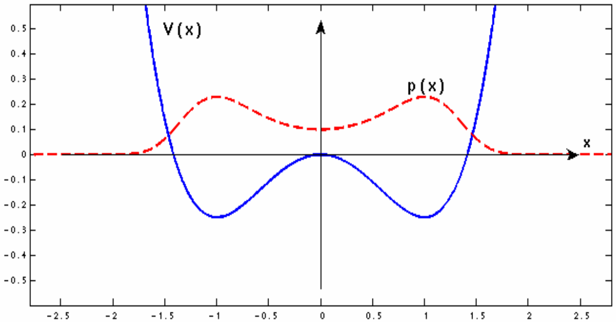


Fig. 2 *V(x)* and its corresponding probability density function *p(x)*.

Another way to estimate a probability density function is by a histogram (Eubank & Farmer 1997). Assuming a system described by $\dot{x} = f(x) + Noise$, we can use the computer to numerically simulate the asymptotic behavior of parameter x and thus approximate the solutions to our equation (Higham 2001). We then partition the state space of x (the x axis) into a number of bins, and we count the number of solutions falling in each bin. An estimate of the density measure of each bin is n^i / N, where n^i is the number of solutions falling in the bin and N is the total number of points. A histogram estimation of a density function is shown in Fig. 3. At the left panel, we show the potential V(x) with two attractors. This is a symmetric potential in that the wells of the attractors are equally deep. We see that the probability density estimate on the right is also symmetric with roughly equally populated regions around the attractors.

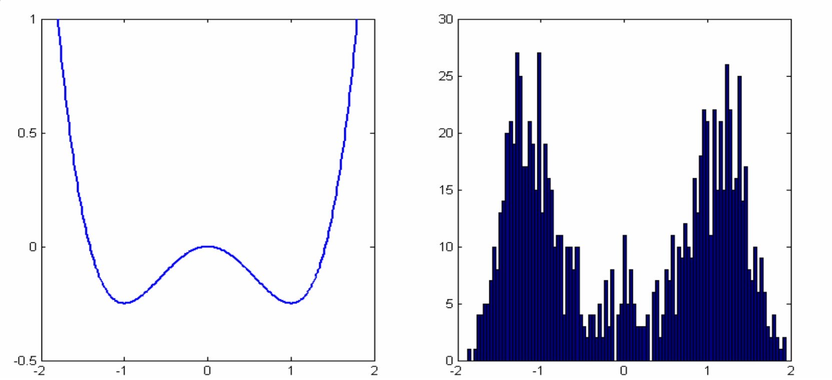


Fig. 3 Symmetric *V(x)* and estimate of density function using a histogram.

The simulation shown in Fig. 4 estimates the probability density function for an asymmetric potential. This allows us to illustrate the notion of the strength of an attractor, which depends on the depth of the minimum in the potential and the steepness of the slope toward the minimum value. Since the attractors are not equally strong, the estimate of the density shows two unequal peaks. The histograms in these figures are generated for a total of $N = 1000$ solution points of $\dot{x} = f(x) + Noise$, where Noise is Gaussian and the initial position of the particle is a random number drawn from the interval $x \in [-1, 1]$.

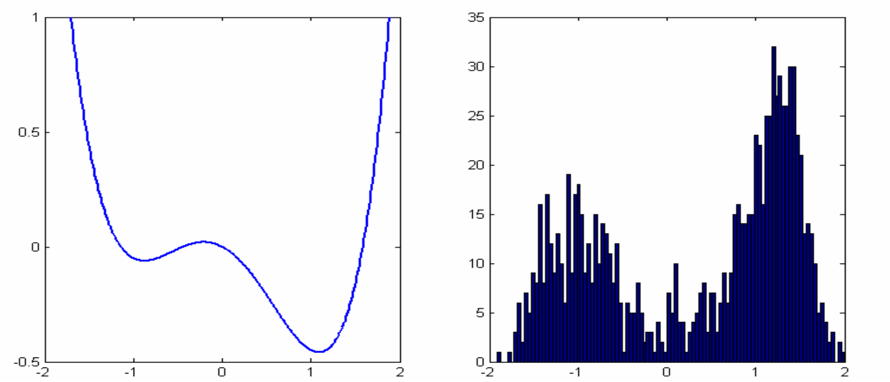


Fig. 4 Asymmetric *V(x)* and estimate of density function using a histogram.

The theoretical results and numerical simulations above show that the preferred regions of order parameters, the attractors, are resistant to noise in a probabilistic sense. It is also true, however, that in behavioral systems this stability in the presence of noise coexists with the flexibility to change. Qualitative changes in a system's order parameter induced by scaling of some 'control' parameter provide a particularly informative entry point into the construction of dynamical models. At a formal level, the ability to change in qualitative ways requires that we relax the property of dynamic stability. Specifically, order parameters must be resistant to noise relative to certain ranges of control parameter values. As the control parameter is scaled beyond some region of values, the order parameter may change abruptly. We are thus led to the essential notion of *nonlinearity*.

A system exhibits nonlinearity when large or discontinuous changes can be observed in the behavior of that system as some control parameter varies smoothly. In a prototypical example of this situation from speech, Stevens (1972, 1989) has argued that the relation between articulatory parameters and their acoustic/auditory output is *quantal* in the following sense. There are certain ranges of articulatory parameter variation within which the acoustic output remains relatively stable. In other ranges, however, small variations in the articulatory parameter cause large (nonlinear) changes in the quality of the acoustic output. Put differently, gradual changes in some articulatory parameters lead to qualitatively distinct acoustic outputs. In another example from biological coordination, Kelso (1984) observed that when adults are asked to move their index fingers in an anti-phase pattern (both fingers move to the left or the right at the same time), they can perform this task over a wide range of cycling frequencies. But as frequency is increased, subjects show a spontaneous shift to an in-phase pattern, that is, to a pattern where the fingers move toward each other or away from each other at the same time. The important point in these examples is that scaling of a continuous parameter results in qualitative changes, the shift from one stable mode to another. Such qualitative changes are commonly referred to as bifurcations by mathematicians or phase transitions by physicists.

To express nonlinear relations between order and control parameters, we augment the general form of a dynamical system $\dot{x} = f(x) + Noise = -dV(x)/dx + Noise$ with a control parameter *P*. This gives us the equation $\dot{x} = f(x,P) + Noise = -dV(x)/dx + Noise$. In general, as *P* changes continuously, the corresponding solutions to our equation also change continuously. But, when *P* crosses a critical value the system may change qualitatively or discontinuously. We can illustrate this fundamental property with a simple mathematical example. Consider a force function parameterized by a control parameter *k* and specified by $f(x,k) = -kx - x^3$. We are interested in what happens to *x* – ultimately the solutions to our equation – as the control parameter *k* is varied. The potential corresponding to our force function $V(x) = kx^2/2 + x^4/4 + (C)$ is plotted in Fig. 5 for various values of *k*. For $k < 0$, as shown in the top row of Fig. 5, the control parameter changes but the system retains a qualitative sameness of form. The two minima in the valleys of *V(x)* represent the stable fixed points of *x*. These are the attractors, the preferred regions within the continuum of *x* where the particle ends up. But as *k* passes through zero, suddenly a qualitative change occurs. The system changes to a monostable regime, showing a single attractor, and retains this form for different values of the control parameter.

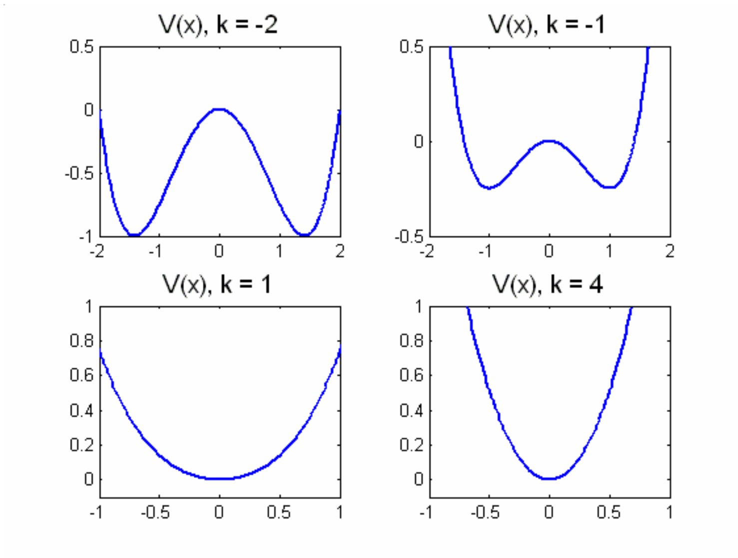


Fig. 5 Potential as a function of a control variable k.

Within the ranges $k > 0$ or $k < 0$, variation in k does affect smoothly the attractor landscape (this is known as scaling). For example, from $k = 4$ to $k = 1$ there is a change in the stability of the attractor, as is evident from the flattening of the walls in the attractor's basin and hence the flattening of the corresponding probability density function. But as long as *k* does not pass the critical value of *0*, that is, within the ranges $k > 0$ or $k < 0$, the macroscopic form of the system remains unaltered. It is only when *k* passes a certain critical value (here, $k = 0$) that a qualitative change takes place.

Next, we consider how the dynamical concepts introduced here can be applied to modeling speech patterns. Section 3 discusses dynamics in the domain of linguistic representations. Section 4 addresses the incompleteness of phonological neutralization, and section 5 focuses on the relation between continuous and discrete dimensions of speech patterns in vowel harmony.

3 Dynamics in Representations

The main claim of this paper is that phonological cognition has a dynamical basis. This claim finds a precursor in a dynamically-based theory of phonological representations developed by Browman & Goldstein (1986 *et seq*.). To highlight connections between our claim and previous work, we provide a synopsis of the main aspects of that theory. The discussion proceeds by contrasting the dynamically-based theory of representations to the view in the background of most work in phonological theory.

The dominant view of phonological representations was established with the development of generative phonology (Chomsky & Halle 1968) and was subsequently elaborated and refined in important ways (Goldsmith 1976, Clements 1985, Sagey 1986). According to this view phonological representations consist of discrete, mental 'sound images' abstracted away from the continuity of phonetic substance. The mental representation of a word consists of a linear sequence of segments along with a set of features characterizing each segment. For example, the word *map* is represented as the sequence of segments [m-ɑ-p] along with the feature bundle for each segment as shown in (2). These features encode linguistically significant distinctions among sounds, that is, they provide a basis of an attested contrast within some language or they play a role in formulating distinctions among the phonological systems of different languages (Anderson 1974).

(2) Phonological representation of *map*

$$\begin{bmatrix} -\text{syllabic} \\ -\text{cont} \\ +\text{labial} \\ +\text{nasal} \\ \cdots \end{bmatrix} \begin{bmatrix} +\text{syllabic} \\ +\text{low} \\ +\text{back} \\ -\text{nasal} \\ \cdots \end{bmatrix} \begin{bmatrix} -\text{syllabic} \\ -\text{cont} \\ +\text{labial} \\ -\text{nasal} \\ \cdots \end{bmatrix}$$

The relation between the discrete symbols of this representation and the continuous phonetics of speech is to be fleshed out by a process of translation from discrete units to continuous physical properties of an articulatory and acoustic nature (Keating 1988, 1990, Cohn 1990, Coleman 1992).

A relatively more recent and less widely explored view of phonological representations is fleshed out in Browman & Goldstein's (1986, 1989, 1990, 1992, 1995) research program. The central hypothesis of this program is that the variable and context-specific movements of vocal tract activity can be captured by underlying invariant dynamical units of action called *gestures*. Moreover, the hypothesis goes, gestures can provide a cognitively plausible basis for a theory of phonological representations. A gesture is a dynamically defined spatio-temporal unit consisting of the attainment of a constriction at some location in the vocal tract. A gesture's spatial dimensions consist of the *articulator set* employed in producing the constriction, and the *vocal tract variables* of constriction location (CL) and constriction degree (CD). For example, the consonants [p], [b], [m] share a lip closing gesture effected by a coordinated activity of the articulator set of the upper lip, lower lip and jaw, with CL = {labial} and CD = {closure}. Gestures contrast on the basis of their tract variables: [b]'s CL is {labial} whereas [d]'s is {alveolar}, [d]'s CD is {closure} whereas [s]'s is {critical}, the constriction degree value of fricative consonants, and so on. Descriptors like {labial}, {alveolar} are cover symbols for ranges for language-particular numerical values corresponding to different points of articulation along the longitudinal axis of the vocal tract. The correspondence between gestures and the segments or features of symbolic theories is not one-to-one. In general, a segment is a constellation of gestures. For example, [p] consists of a lip closing and a glottal abduction gesture, [m] consists of a lip closing and a velic opening gesture, and [t] consists of a tongue tip closure at the alveolar place and a glottal abduction gesture. Some features correspond to gestures (e.g. [+nasal] corresponds to a velic opening gesture), but other features correspond to values of tract variables (e.g. [+labial] corresponds to CL {labial}).

We now turn to the most relevant aspect of gestures, namely, that they are dynamical units. There are two senses of 'dynamical' here. One sense implies movement. As usually described, gestures do involve movement. However, the claim that movement is a definitional property of gestures is potentially confusing: How can entities defined in terms of physical movement serve as mental units of representation in phonological cognition? To avoid confusion we must turn to the other sense of dynamical, which is also more central to the point we want to emphasize here. Gestures are dynamical in that their mathematical model employs concepts from nonlinear dynamics. As a gesture unfolds, its tract variables CL and CD change in value over time and can thus be modeled as dynamical systems. Take, for example, an alveolar stop /d/ with CD = {closure}, CL = {alveolar}. The CD tract variable, independent of its initial value, that is, independent of the constriction degree the tongue tip happens to have before the onset of movement for [d], will approach its target value which is {closure} or 0 mm (in fact, 3.5 mm is used to model surface compression). It is this continuous change in the value of CD that is modeled as a dynamical system. The same applies to modeling CL. The specific mathematical model employs the point-attractor dynamics of damped mass-spring systems. These are governed by the equation in (3), where x_0 is the equilibrium position or the spring's rest length (corresponding to the tract variable value), k is the spring's stiffness (determining the time to target or duration of movement), and b corresponds to the damping (which is set to critical, so that the system approximates the target without oscillation).

(3) Dynamical model of gestures: $d^2x/d^2t + bdx/dt + k(x - x_0) = 0$

It is useful to consider how this dynamical model provides a basis for capturing the variability and context-specificity of speech movements. Like any differential equation, the equation above describes how tract variable x changes over time. Different evolution paths of x values can be obtained by tuning the system parameters, that is, the initial position of x, the stiffness k, and the equilibrium position x_0. For example, consider the fact that the kinematics of a lip closure in [aba] vs. [ibi] are different in that the lips must travel over a greater distance in attaining lip closure in [aba] than in [ibi]. This is because the jaw position for the contextual vowels is lower in [aba] than in [ibi]. Since the jaw is physically linked to the lower lip, when a [b] gesture becomes active the position of the lower lip will be different in the two conditions; hence, the different trajectories of lip movement. Nevertheless, the invariance of a [b], that is, lip closure is attained as a consequence of the dynamical model. Similarly, the kinematics of a [b] are different under different rate conditions and this corresponds to different stiffness values in (3). In effect, then, this equation encodes a family of trajectories sharing the task specified by the target CL, CD values. Gestures in this respect are like other targeting movements in living systems with the intrinsic property of being self-equilibrating. Indeed, the mathematical model for gestures derives from a general theory of action and is not in its underlying principles particular to speech (Saltzman 1995).

One central idea about the role of dynamics emerges from this discussion. The dynamical model states an intimate and immediate link between the abstract invariance of phonological entities ([b] vs. [d]) and the continuity of their phonetic substance. Gestures can be said to be abstract and invariant in the following respects. First, gestures are specified by their tract variables, CL and CD. These values are constant and do not change during the lifetime of a given gesture. For example, the gesture for [b] specifies a different constriction degree from that of [w] and a different constriction location from that of [g], encoded in terms of two distinct values along the longitudinal axis of the vocal tract. Gestures are invariant in that their model does not change during their lifetime. The equation in (3) expresses an invariant underlying law which gives rise to a gesture's kinematic patterns. The kinematics themselves are continuous and context specific. But the underlying law governing this surface variability is an abstract, invariant model, that is, a mathematical statement in the form of a differential equation. Equivalently, the actual trajectory for a gesture is not encoded time point by time point. Rather, the continuous trajectory unfolds from the dynamical model of gestures.

To sum up, a distinctive property of the dynamical view of phonological representations is its use of a mathematical language in which continuity and discreteness coexist. Our aim in the next two sections is to extend the argument for the dynamical basis of phonological cognition from representations to organizational principles of linguistic grammars.

4 Interaction between grammar and communicative context

The translational view of the relation between phonology and phonetics expresses a valid intuition, namely, that there is a distinction to be made between continuous and discrete aspects of phonological cognition. For certain phenomena, however, the specific way of drawing that distinction within the translational approach proves to be too rigid. This section takes up one class of such phenomena whose defining property is the interaction between discrete grammatical requirements and continuous communicative or environmental variables. As we will see, the translational view precludes this kind of interaction. In turn, this interaction underscores the need for a formal language integrating discrete and continuous aspects of the phonology-phonetics system. Using nonlinear dynamics, our aim is to provide a way of reconciling the valid intuition of the translational view with cases of interactionism.

A basic property of phonological systems is that the phonetic properties of sounds are not equally distributed across all syllabic positions. For example, German is well known for its devoicing of consonants in syllable-final position. Thus, *Rad* 'wheel' is produced as [ʁat] ‘wheel-nominative’, with a voiceless [t], when that consonant is at the end of a syllable, but with a voiced [d] in all other contexts, e.g. [ʁadəs] ‘wheel-genitive’. Moreover, whereas the produced form of ‘wheel’ alternates between a voiced and a voiceless final consonant, other words do not show this alternation. For example, *Rat* ‘advice’ is always produced with a [t], [ʁat] ‘advice-nominative’, [ʁatəs] ‘advice-genitive’. Because the contrast between ‘wheel’ and ‘advice’ is neutralized in some positions, in that both are produced with a final [t], this phenomenon is referred to as (*contrast*) *neutralization* (Bloomfield 1933:218, Trubetzkoy 1969:213).

When the performance of German speakers is closely observed, however, it deviates slightly from the simple description of neutralization above. The basic result is that the [t] in [ʁat] ‘wheel-nominative’ is pronounced quantitatively differently from the final consonant of *Rat* ‘advice’. Specifically, in [ʁat] ‘wheel-nominative’, the [t] shows traces of voicing which make it distinguishable from the completely voiceless [t] of [ʁat] ‘advice-nominative’. Moreover, this difference depends on the communicative context. Port & Crawford (1989) have shown this by setting up different experimental conditions in an attempt to study speakers’ behavior in more or less naturalistic communicative contexts. In one condition, speakers read a given word-list. In another condition, speakers are asked to read sentences like *Ich habe Rat(Rad) gesagt; nicht Rad(Rat)* “I said Rat(/Rad) not Rad(/Rat)” while a German assistant, present in the experimental setting, is assigned the task of writing down the order of the test words in such sentences. In the latter task, speakers are encouraged by the context to convey the difference between *Rad* ‘wheel’ and *Rat* ‘advice’ more than in the word-list reading task. The result is that the ‘neutralized’ final consonant of *Rad* in [ʁat] ‘wheel-nominative’ shifts more toward a voiced [d] than in the word list reading task.[1]

We now turn to see how neutralization has been analyzed in phonological theories. We will consider two kinds of analyses, a standard one based on rules (Chomsky & Halle 1968) and an alternative based on constraints (Prince & Smolensky 1993/2004, 1997). Both of these turn out to be unable to capture the subtle phonetic differences in neutralization. But a dynamical formulation of the constraint-based analysis will provide us with an entry to our proposed model, which integrates the discreteness of the phonological patterns with the flexibility of phonetic performance.

The standard way of expressing the German neutralization phenomenon in linguistic grammar is to say that German speakers have internalized a rule which turns voiced consonants to their voiceless counterparts when they are in a syllable-final position. This is the rule of Final Devoicing shown in (4). In addition, German speakers store distinct representations for the basic forms of 'wheel' and 'advice', /ʁad/ and /ʁat/ respectively, in their mental lexicon. Phonological rules, like Final Devoicing, mediate between these input or underlying representations and their surface realization. Following standard practice, we use '/ /' for the representations in the mental lexicon and '[ ]' for their surface representation. Thus, Final Devoicing applies to the mental lexicon entry for 'wheel', /ʁad/, to produce [ʁat] but fails to apply to /ʁat/ or /ʁad + əs/ as shown in (4).

(4) Final Devoicing (***FD***)[2]: [+Voiced, –Sonorant] → [–Voiced] / __ $]^{\sigma}$

---

[1] There is an extensive literature on the phonetics of neutralization. See Blumstein (1991) for cogent discussion of neutralization in the context of the phonology, phonetics relation. See Dinnsen (1985) for a review of other instances of incomplete neutralization, Dinnsen & Carles-Luce (1984) on Catalan final devoicing, Fougeron & Steriade (1997) on French schwa elision, Charles-Luce (1993, 1997) on Catalan voicing assimilation and English flapping, and Piroth & Janker (2004) for some recent results on German neutralization. See Jongman (2004) and Kim & Jongman (1996) for a case of complete neutralization in Korean.

[2] In the rule-based theory of phonological patterns, rules have three parts (Chomsky & Halle 1968). The focus part '[+Voiced, –Sonorant]' lists the features specifying the class of segments undergoing the rule (here, the voiced

Sample phonological derivations:

	Input		*Output*
a.	/ʁad/	→	[ʁat] (***FD*** applies)
b.	/ʁat/	→	[ʁat] (***FD*** can't apply as /t/ is not [+Voiced])
c.	/ʁad + əs/	→	[ʁadəs] (***FD*** can't apply as /d/ is not syllable-final)

Optimality Theory (OT), as developed in Prince & Smolensky's work, presents an alternative conception of grammar where the output of phonological computation is the result of constraint optimization as opposed to rule-based operations. Two broad classes of OT constraints are recognized. Markedness constraints define dimensions of harmony in linguistic structures, such as "a syllable must have an onset" or "a vowel must be oral (not nasalized)". Faithfulness constraints require that the output of phonological computation preserves properties of its input in the mental lexicon. A fundamental tenet of OT is that phonological patterns are the result of constraint optimization. To exemplify with voicing neutralization, there are two relevant constraints. The MARKEDNESS constraint requires coda consonants to be voiceless. The FAITHFULNESS constraint requires identity between the representation of a consonant in the mental lexicon and its surface form. When the input is /ʁad/, the constraints are in conflict. Output [ʁad] preserves the voicing as demanded by FAITHFULNESS but violates MARKEDNESS. Output [ʁat] satisfies MARKEDNESS but violates FAITHFULNESS. Since in German the surface form is [ʁat] the constraints are prioritized so that MARKEDNESS is ranked higher than FAITHFULNESS. This is written as MARKEDNESS >> FAITHFULNESS. When the constraints are

---

obstruents), the change part '→ [–Voiced]', specifies the featural change effected by the rule, and the environment part '/ __]$^{\sigma}$' specifies the context in which the rule applies (here, at the ends of syllables).

ranked differently, FAITHFULNESS >> MARKEDNESS, the result is preservation of the underlying voicing as is the case in English. In OT, then, neutralization phenomena are an outcome of constraint optimization under the ranking schema MARKEDNESS >> FAITHFULNESS. Intuitively, the relative ranking of FAITHFULNESS with respect to MARKEDNESS results in the loss of contrast in certain positions, e.g. the loss of the voicing contrast in the coda position of /ʁat/ vs. /ʁad/ since for both of these inputs the output of the grammatical computation is [ʁat].

We now extract two shared properties of the above analyses which turn out to be inconsistent with the phonetic facts of neutralization. One is that grammatical computation is based on discrete categories – the representations and the rules or constraints referring to these are spelled out in the language of symbols. The variable phonetic outputs in the neutralization facts cannot be fully described using categories. The final [t] of [ʁat] 'wheel-nominative' is voiceless but it is quantitatively different from other voiceless segments, e.g. [t] of [ʁat] 'advice-nominative'. The other assumption is that phonological computation precedes the phonetic implementation of grammatical outputs. Recall that in the output of phonological computation, whether rule- or constraint-based, the voicing value of the final consonant of *Rad* 'wheel-nominative' is [–Voiced]. This eliminates the contrast between the final consonants of *Rad-Rat* at the output of phonological computation, exactly as a neutralizing rule or ranking schema should do. Consequently, the phonetic implementation transducer, whose role is to flesh out the phonological output as vocal tract action, is now unable to deliver the difference observed in the realization of the final consonants in *Rad* vs. *Rat*.

We thus revisit the neutralization problem with the goal of accounting for the fact that the speakers phonetic outputs adapt flexibly in a purposive way to demands dictated by the communicative context. This requires an explicit link between context and grammar so that the two can interact in ways that derive the observed systematicities. Using the OT analysis as an entry point, our plan is to formulate sub-symbolic, dynamical versions of the main forces at work here. Specifically, the faithfulness and markedness constraints will be situated in the continuous realm of 'forces' on phonetic variables as opposed to symbols. This will provide a basis for capturing the systematic scaling of the phonetic output as contextual parameters change.

We begin by a dynamical formulation of the faithfulness constraint. Consider a communicative act wherein the speaker's goal is to convey the word *Rad*. Since in the mental lexicon *Rad* is stored with a voiced final segment, /ʁad/, we can express the intention to convey this form by an intentional dynamics which contributes an attractor at the required value of voicing. The same applies to *Rat*, where the intention is voiceless. This idea is depicted in Fig. 6 showing the potential functions for voiced and voiceless intentions. The x axis represents the space of all possible voicing values, which we assume can be indexed by the degree of glottal aperture, one of the tract variables in the gestural model discussed in section 3. The voiceless intention is represented with an attractor, a minimum of the potential function, at some positive value of glottal opening x_0 and the voiced intention with an attractor at the some negative value of glottal opening, $-x_0$ (the actual numeric values are not crucial in the present context).

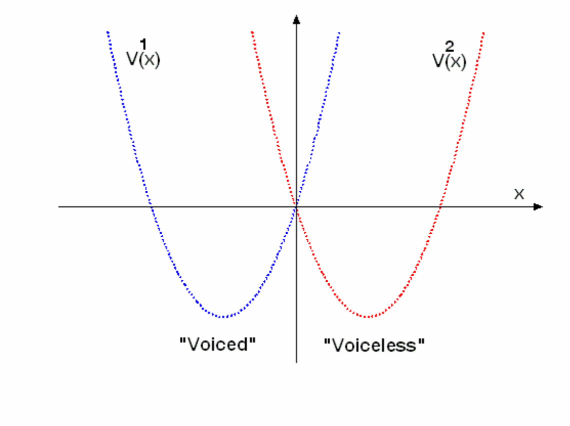


Fig. 6 Potential functions for voiced and voiceless intentions.

We can visualize the glottal aperture variable as the position of a particle left in one of the potentials above. Intention acts as a force constraining the glottal aperture variable to fall within a region of values around the mean characteristic for the intended voiced/voiceless consonant. Hence, the attractors at x_0, $-x_0$ describe qualitatively distinct states of the voicing system or, in other words, they describe a dimension of order in the continuous space of phonetic variables. In a symbolic statement of the voicing contrast, the situation is written in terms of the mental binary distinction [±Voiced]. This notation describes two discrete categories abstracted away from phonetic substance but ultimately translated to that substance via a phonetic implementation transducer. The difference between the symbolic and the dynamical formulation is that in the latter (but crucially not in the former) the distinct modes of voicing are inseparably linked with the phonetic substance. These modes are not derivationally antecedent to that substance and therefore they do not need to be translated to that substance. Eco, who has studied the foundational notion of symbol closely, writes: “One cannot speak of a form without presupposing a matter and linking it immediately (neither before nor after) to substance” (1984:23).

Formally, as in any (autonomous) dynamical system, intention is modeled by a differential equation of the form $\dot{x} = I(x)$. $I(x)$ is the simplest function that admits a stable fixed point at the (intentionally) required value of voicing, or $I(x) = \theta(x^{REQ} - x)$. In this function, θ is a linear term representing the relative strength of the intentional contribution and can be ignored for now. The term x^{REQ} takes values from $\{-x_0, x_0\}$, that is, the glottal aperture values corresponding to [+Voiced], [–Voiced] consonants. To derive the intention potential, we use $\dot{x} = I(x) = \theta(x^{REQ} - x) = -dV(x)/dx$, and by basic calculus we can derive $V(x) = \theta x^2/2 - \theta x^{REQ} x$, up to a constant C which can be dropped since it is of no qualitative significance in the context of this discussion and the simulations. The potentials for two values of x^{REQ}, $\{-x_0, x_0\}$, are shown in Fig. 6 above.

Next, we describe the continuous equivalent of the markedness constraint, requiring coda consonants to be voiceless. To state this constraint in dynamical terms, we specify a potential function that contributes an attractor at the appropriate value of voicing. The required potential is shown in Fig. 7. Since codas can only be voiceless, the attractor is at a value characteristic of voiceless consonants. A particle left in this potential ends up at the minimum representing a value of glottal aperture appropriate for voicelessness.

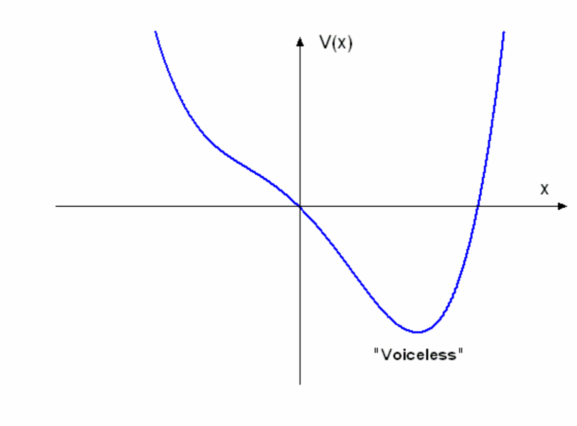


Fig. 7 Dynamical model of markedness constraint.

Formally, the markedness constraint can be defined by a differential equation of the general form $\dot{x} = M(x)$. As a working hypothesis, we assume that $M(x) = \dot{x} = -k + x - x^3$. Given that $-dV(x)/dx = \dot{x}$, we can compute by integration the potential $V(x) = kx - x^2/2 + x^4/4$. This V(x) is the potential shown in Fig. 7 with an attractor at some positive value of glottal opening characteristic of voiceless obstruents.

We can now look at the interaction between the markedness and faithfulness forces in the dynamical setting. Formally, the simplest model for this interaction corresponds to a linear combination of the M, F forces, that is, $dx/dt = M(x) + F(x)$. Consider what this model predicts for the intention to convey *Rad* ‘wheel’. In German, the *M* force dominates the *F* force. Thus, when the two forces combine, the result is pulled toward the voiceless attractor. But the presence of the dominated *F* force is nonetheless felt. Since intention contributes an attractor at the intended [d] form, the attractor for the combined system is a little closer to that of [d] than in the case when the intention is *Rat* 'advice'. This competition between the *M*, *F* forces is schematized below for various values of intentional degree. This is the scalar control parameter θ in the dynamics for intention $I(x) = \theta(x^{REQ} - x)$ representing the relative strength of the intentional contribution. A value of intent close to 0 corresponds to a context where the speaker's intention to communicate the contrast between *Rat* and *Rad* is weak, as would be the case in the word-list reading, assistant-absent condition. Higher values correspond to communicative contexts with stronger requirements for expressing the contrast as would be the case in the assistant-present condition. It is observed that, as intentional strength increases, the potential is gradually scaled so that the attractor drifts toward the left or toward more voicing.

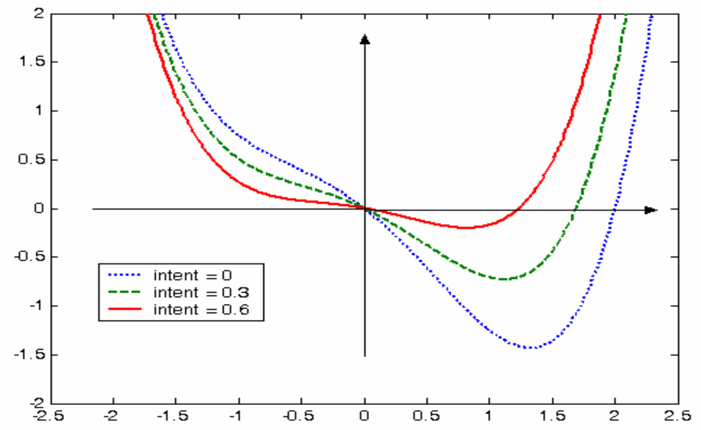


Fig. 8 Scaling of phonetic output for different intentional degrees.

The drift in the attractor of the phonetic output above is a reflex of constraint conflict between the *M*, *F* forces relative to different contextual conditions. This drift is beyond the scope of the standard OT model where constraint conflict is at the symbolic level. When constraint interaction is trapped at the level of symbols, it can only yield the same categorical representation [ʁat] for *Rad*, *Rat*.

The dynamical formulation of constraint conflict maintains the discreteness of the symbolic formulation while embedding it in a continuous realm. The discreteness inheres in the attractors of the dynamical system. At the same time, these attractors are embedded in continuous phonetic dimensions. Consequently the attractors can be modulated by variation in context in a way that derives the subtle differences in *Rad* vs. *Rat* and their dependency on the communicative context. At a more broad level, the main proposal here is that it is both necessary and promising to do away with the metaphor of precedence between the qualitative phonology and the quantitative phonetics, without losing sight of the essential distinction between the two. This requires a coherent way to make discreteness and continuity coexist within the same formal language. The mathematics of nonlinear dynamics satisfies this requirement.

5 Dynamical links between categorical alternations and continuous phonetics

We now turn to the problem of how low-level spatial phonetic properties of vowels relate to the high-level phonological behavior of suffix choice in the phenomenon of vowel harmony. Though this problem derives from a rather different phonological phenomenon from that of consonant devoicing, discussed in the previous section, the two share a common formal challenge, namely, relating continuous phonetic variables to discrete phonological patterns. As in the previous section, we present a dynamical model which allows us to formalize the link between the discreteness of phonological form and the continuity of phonetic substance in which that form is embedded.

5.1 Vowel harmony and transparency

Vowel harmony is a systematic regularity found in many languages by which vowels in a word must agree in terms of certain phonetic properties. For example, vowels in Hungarian words tend to be drawn either from the set of vowels /i í e é ö ő ü ű/ articulated with a frontward movement of the tongue body and transcribed in IPA as [i iː ɛ eː ø øː y yː], or from the set of vowels /u ú o ó a á/ articulated with a backward movement of the tongue body and transcribed as [u uː o o ɔ ɑː] (Siptár & Törkenczy 2000). In terms of phonological features, the 'front' vowels share the feature [–back], and the 'back' vowels share the feature [+back]. In the description of harmony patterns we will use Hungarian orthography where the acute accent denotes length, and the umlaut denotes front round vowels.

The phonological consequences of vowel harmony are most readily observed in suffix vowel alternations where the [±back] quality of the suffix vowel is determined by the [±back] quality of the stem vowel. For example, the Dative suffix alternates between two forms, one with the front /e/ and another with the back /a/, as a function of the stem vowel: *ház-nak* 'house-Dative' but *kéz-nek* 'hand-Dative'. Because it determines the suffix form, the first stem vowel is called the trigger and the suffix vowel is called the target of the harmony pattern.

It has been proposed that a natural basis for vowel harmony can be traced to the low-level phonetic effects among vowels in consecutive syllables (Fowler 1983, Ohala 1994). The crucial fact is that vowels exert influences on neighboring vowels across intervening consonants, the so-called *V-to-V coarticulation* (Öhman 1966). However, V-to-V coarticulation is a quantitative pattern whose degree varies depending on the quality of intervening consonants, stress distribution, and other factors (e.g. Recasens 1999). It thus remains to be shown how such variable and quantitative coarticulation effects are to be linked to the binary [±back] character of suffix alternations. As emphasized all along, addressing this challenge requires the appropriate formal tools enabling us to express relations between discrete and continuous aspects of complex systems.

A second problem with the proposed phonetic basis of vowel harmony is that many languages with harmony include vowels that disagree with their adjacent vowels. These vowels are called *transparent* because their most well-known property is that they may intervene between the trigger and the target vowel even when they bear the opposite value for the harmonizing feature. For example, in Hungarian, *papír* selects [+back] suffixes, such as *nak* ‘Dative’, *ház* ‘Allative’*, tól* ‘Ablative’*, ban* ‘Inessive’, in agreement with the [+back] value of the initial stem vowel and despite the intervening [–back] value /í/. The rest of the Hungarian front unround vowels, that is, /i, é, e/, behave similarly. Here are some representative examples of words with these vowels: *gumi-nak* ‘rubber-Dative’, *kávé-nak* ‘coffee-Dative’, and *hárem-nak* ‘harem-Dative’. Transparent vowels, then, present a challenge to the proposal that vowel harmony has its basis in V-to-V coarticulation effects between consecutive vowels. This is because the [+back] quality of the suffixes in the words above cannot plausibly be derived via V-to-V coarticulation when their preceding vowels are specified as [–back].

At the heart of the problem that transparent vowels pose for the phonetic basis of vowel harmony is an implicit assumption about their representation. This assumption is that the phonological category of a transparent vowel is invariant across different contexts and irrelevant to the quality of the suffix following the transparent vowel. In an impressionistic sense, the transparent vowels in words like *bul**i**-nak* ‘party-Dative’, *h**í**d-nak* ‘bridge-Dative’ or *mam**i**cs**i**-nak* ‘mother-Diminutive-Dative’ are not perceptually different from those in *bil**i**-nek* ‘pot-Dative’ or *v**í**z-nek* ‘water-Dative’. Hence, they are assumed to be invariant across these different contexts. However, it is well-known that for vowels a relatively stable acoustic output can be produced using multiple articulatory strategies and constriction locations. Independent work by

Stevens (1972, 1989) using simple tubes, and Wood (1979) using natural human vocal tract profiles, has shown that the acoustic outputs for non-low front vowels—exactly the transparent vowels of languages like Hungarian and Finnish—are insensitive to a limited amount of variation in the horizontal position of the tongue body. Therefore, the impressionistic perceptual invariance of transparent vowels in different contexts does not necessarily imply their articulatory invariance.

Our hypothesis, pursued in experimental work (Benus *et al*. 2004, Gafos & Benus 2003), is that transparency emerges from non-linearities in the relation between articulation and sound. In a nutshell, we hypothesize that the /í/ in *zafír-ban* 'sapphire-Inessive' is retracted articulatorily as compared to /í/ in *zefír-ben* 'zephyr-Inessive', but that this retraction falls within that limited region of articulatory variation that does not result in any significant acoustic consequences. If this hypothesis is correct it would provide a basis for a principled understanding of the co-occurrence of two properties of the phenomenon, the nature of the harmonizing parameter (tongue body retraction) and the set of transparent vowels in Hungarian (/í, i, é, e/).

The hypothesis that transparent vowels do participate in vowel harmony by sub-categorical changes in their tongue body position may help understand other heretofore recalcitrant generalizations in the Hungarian vowel harmony system (Benus 2005). First, stems with only transparent vowels, henceforth T stems, may trigger both front and back suffixes (Vago 1980). The majority of T stems trigger front suffixes (*cím-nek* 'address-Dative', *szél-nek* 'wind-Dative'), but approximately sixty T stems trigger back suffixes (*síp-nak* 'whistle-Dative', *cél-nak* 'aim-Dative'). From the perspective of categorical representations, this situation is paradoxical. Since the vowels in *szél*, *cél* are represented identically, they are expected to select the same suffix, but they do not. However, if there are systematic phonetic differences in tongue body position between these two groups of stems, they may provide a basis for their divergent suffix choices.

Second, the height of the transparent vowels in stems where they are preceded by a back vowel (BT stems) affects the choice of the suffix (van der Hulst 1988, Hayes 2004). The lower the vowel, the more likely it is that the suffix following BT stems is front. For example, Hungarian speakers accept only back suffixes after *papír* but allow free variation between a front and a back suffix after *hotel*. This latter pattern is known as *vacillation* in the Hungarian literature. What is the difference between /í/ and /e/ that may be responsible for their distinct suffix selection patterns? An answer might come from differences in the articulatory-acoustic properties of high vowels such as /í/ versus lower vowels such as /e/. Specifically, the lower and more retracted tongue body of /e/, as seen in our experiments, reduces the region of acoustic insensitivity to articulatory variation, thus predicting that /e/ cannot be retracted to a degree comparable to that of /í/.

Finally, increasing the number of transparent vowels following a back vowel decreases the likelihood of selecting back suffixes. For example, *mam-i*, *mam-csi*, both meaning 'mother-Diminutive', select back suffixes as in *mam-i-nak*, *mam-csi-nak* 'mother-Diminutive-Dative'. However, when *–i*, *–csi* are combined as in *mam-i-csi*, both front and back suffixes become acceptable, *mamicsi-nak* and *mamicsi-nek*. Compare also *kabin-nak* 'cabin-Dative' vs. *aszpirin-nak/nek* 'aspirin-Dative' and *Acél-nak* 'Acél-Dative' vs. *Acélék-nek/nak* 'Acél-Collective-Dative' (Farkas & Beddor 1987, Ringen & Kontra 1989, Kaun 1995, Hayes 2004). Under the assumption that all vowels, including the transparent ones, participate in vowel harmony, the position of tongue body for the stem-final vowels in stems like *mami* is predicted to be different from the position of the stem-final vowel in the stems like *mamicsi*. The vowel /i/ in *mami* is less advanced or more retracted than the stem-final /i/ in *mamicsi* because the additional front vowel in the second syllable of *mamicsi* eliminates partially the influence of the initial back vowel on the final /i/.

In short, our hypothesis predicts that transparent vowels show systematic articulatory differences in different contexts and that these differences are linked to the harmonic behavior of these vowels. In the next section, we present the results of our experiments on the articulatory properties of transparent vowels.

5.2 Articulatory characteristics of transparent vowels

To examine the articulation of transparent vowels in vowel harmony, we used two experimental techniques, electromagnetic articulometry and ultrasound. Before a presentation of our results, we give a brief overview of these techniques.

In electromagnetic articulometry or EMMA (Perkell *et al*. 1992), an electromagnetic field is used to track movements of small receiver coils attached to the speech articulators. Three transmitter coils are fixed on a plastic apparatus surrounding the speaker's head. The transmitted coils produce alternating magnetic fields at different frequencies in the range of about 10 KHz. Small receivers, about 2 mm in diameter, are attached on the speech articulators using special adhesive. The electromagnetic field from the transmitter coils passes through the receiver coils and generates an electric signal. The voltage of this signal is inversely related to the distance of the receiver relative to the transmitter coils. This relationship is used to calculate the position of the receivers as a function of time. The voltages in the receiver coils are captured at a sampling rate of 500 Hz. Due to their high sampling rate, EMMA systems are currently the most accurate tools for the collection of real-time lingual articulatory movement data.

A representative recording from the EMMA system is in Fig. 7. This figure shows one instance of a sentence containing the target word *zafírban*. The signals from top to bottom represent the acoustic waveform of the entire sentence, the acoustic waveform of the target word with an approximate segmentation into phonemes, and the vertical (solid curve) and horizontal (dashed curve) position of the receivers attached on the tongue tip (TT), body (TB), dorsum (TD), and upper and lower lips (UL, LL). The axis at the bottom depicts time in milliseconds. As the tongue body smoothly moves from vowel to vowel in the sequence *a-í-a* of *zafírban*, the receivers on the TB, TD can be seen to trace a bell-shaped trajectory from a retracted position for /a/, with low horizontal values of TB and TD, to an advanced position for /í/, with higher horizontal positions for TB and TD, and back to a retracted position of the final /a/. To quantify the spatial properties of transparent vowels, here the /í/ in *zafírban*, we identified the maximal horizontal positions of the TB, TD receivers during that vowel. These are shown by the 'max' labels in Fig. 7, corresponding to the peaks in the horizontal trajectories of TB, TD. At these time points the TB, TD receivers are at their most advanced position.

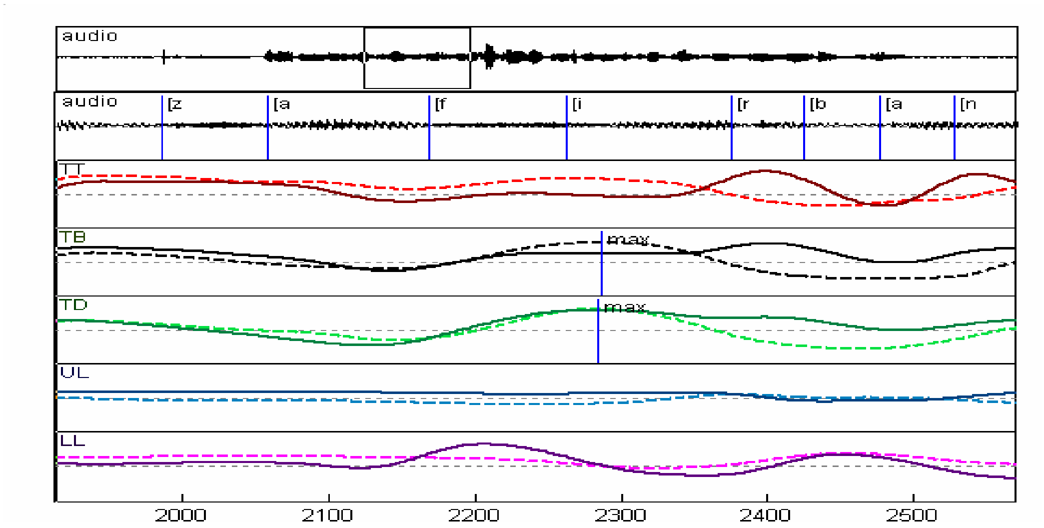


Fig. 7 Articulator kinematics recorded with EMMA.

The other experimental method is ultrasound (Stone 1997). In ultrasound, a probe with a piezoelectric crystal is placed below the subject's chin and emits ultra-high frequency waves. These waves travel through soft tissue and reflect when they reach an interface with a matter of different density such as air. The reflected echo is used to construct a bright white line that shows the boundary between the tongue surface and the air above it. In our experiments, ultrasound images of the tongue were collected at a 30 Hz rate, video recorded, and then digitized. The spatiotemporal resolution of ultrasound is low relative to that of EMMA. The advantage of ultrasound is that it allows visualization of the back region of the tongue's surface from the root to the dorsum. This region is crucially involved in vowel production, but it is usually inaccessible with EMMA due to difficulties with subjects tolerating a receiver attached in that part of the tongue.

To estimate tongue shapes, we first identified the frame with the most advanced tongue position during the production of the transparent vowel (left, Fig. 8). The tongue edge in this frame was then traced using methods of Iskarous (in press) by determining the points of maximal contrast within the selected region (middle, Fig. 8), and fitting multiple snakes into a curve that balances the distance of the points from the curve with the curve's smoothness (right, Fig. 8).

Fig. 8 Tongue surface captured with ultrasound and estimation of tongue contours.

We constructed two sets of stimuli consisting of word pairs where transparent vowels occur in stems triggering front or back suffixes. In the first set, all words were trisyllabic. An example pair from this set is *zefír-ben* 'zephyr-Inessive' / *zafír-ban* 'sapphire-Inessive', where /í/ occurs in *zefír*, triggering a front suffix, and in *zafír*, triggering a back suffix. Such pairs allow us to compare the tongue posture for /í/ in the two vowel harmony contexts, front and back. The second set of stimuli consisted of monosyllabic words. For example, /é/ in *szél* 'wind' was compared to /é/ in *cél* 'aim'. The forms *szél*, *cél* correspond to the Nominative case of the respective nouns, where there is no overt suffix. In other cases with overt suffixes, *szél* triggers a front while *cél* triggers a back suffix: *szél-nek* 'wind-Dative' vs. *cél-nak* 'aim-Dative'. Once again, such pairs allow us to compare the tongue posture for /é/ in the two vowel harmony contexts, front and back. However, they potentially allow us to establish another point. In pairs like *zefír* vs. *zafír*, the difference in suffix choice, front for *zefír* vs. back for *zafír*, is typically ascribed to the presence of a front vs. back stem-initial vowel. In our *szél* vs. *cél* pairs of stimuli, if systematic sub-categorical differences are found in the transparent vowel, then the distinct suffix choices in *szél-nek* vs. *cél-nak* may be related to those sub-categorical differences.

EMMA data from three subjects and ultrasound data from one subject were analyzed. Table (1) shows the data from the EMMA experiments. For the two stimuli sets, trisyllabic (3-syll) and monosyllabic (1-syll), the rows show the mean receiver positions on the tongue tip (TT), body (TB) and dorsum (TD) in the front (F) and back (B) harmony context as well as their Mean Difference (MD = F−B). The absolute value of MD corresponds to the size of the effect, and its sign shows the direction of the effect. If the MD value is positive, the relevant receiver in the back environment is retracted relative to its position in the front environment.[3]

---

[3] The values under F, B are negative because the origin of the coordinate system is approximately at the subject's upper incisors with receiver positions farther inside the mouth represented with progressively decreasing values.

Table (1) Direction and size of the effect of environment in the EMMA data[4]

	Rec.	**ZZ subject**			**BU subject**			**CK subject**		
		F	**B**	**MD**	**F**	**B**	**MD**	**F**	**B**	**MD**
3-syll	**TD**	– 48.02	– 48.97	**0.95****	– 43.12	– 43.51	**0.39****	– 24.59	– 25.58	**0.99***
	TB	– 38.65	– 40.05	**1.40****	– 30.89	– 31.48	**0.59****			
	TT	– 23.41	– 24.73	**1.32****	– 21.68	– 22.07	**0.39****	– 21.83	– 22.08	**0.23**
1-syll	**TD**	– 46.67	– 46.93	**0.26**	– 42.08	– 42.61	**0.53****	– 22.25	– 22.94	**0.69***
	TB	– 36.17	– 36.81	**0.64***	– 29.54	– 30.38	**0.84****			
	TT	– 20.35	– 20.62	**0.27**	– 20.09	– 20.6	**0.51****	– 20.00	– 19.78	**– 0.22**

All MD cells contain positive values except for one cell in CK's TT monosyllabic data. TT is the receiver attached at the most anterior part of the tongue. Thus, its MD values are least relevant to the posture of the tongue body and dorsum, the main determinants of vowel quality (Harshman *et al*. 1977:702). Moreover, for that same subject, the TD value in the monosyllabic data is positive and also shows a significant mean difference between F and B contexts (0.69*). Hence, the main result is that the tongue is more advanced for transparent vowel in stems triggering front suffixes than for transparent vowels in stems triggering back suffixes. This effect is significant for 12 out of 16 MD cells.

The same effect is observed with the ultrasound data. In Fig. 9, we have superimposed tongue postures of /í/, /é/ in a back (dotted lines) vs. front harmony (solid lines) context from subject ZZ. The direction of the effect is the same as that observed with the EMMA data. Transparent vowels in the front context are more advanced than in the back context.

---

[4] ZZ and BU were analyzed with one-way Anovas with the receiver values as the dependent variables and harmonic environment (Front vs. Back) as the dependent variable. Results significant at $p < 0.5$ are shown with '*', and those significant at $p < 0.001$ are shown with '**'. Due to differences in the stimuli sets, the data from the pilot subject CK were analyzed using a paired samples t-test. Also, CK's data do not include values for the TB receiver as technical reasons prevented us from attaching a third coil on that subject's tongue. See Benus (2005) for a detailed description of the results.

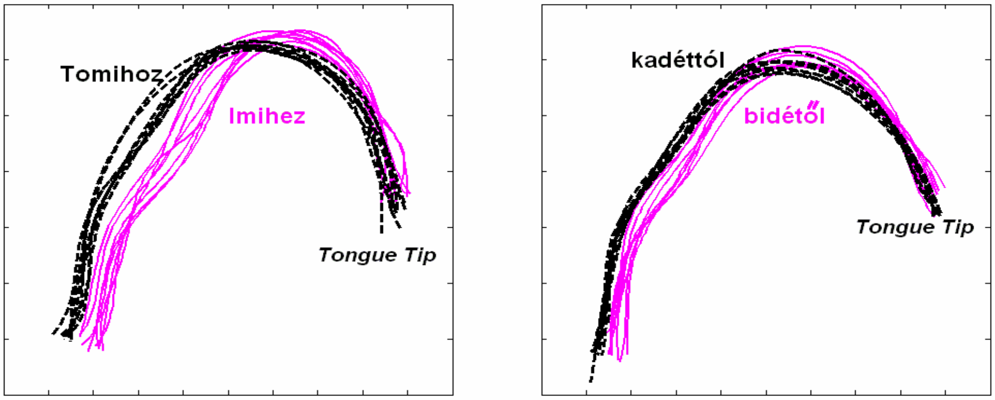


Fig. 9 Effect of environment on /i/ (left) and /é/ (right).

We observed differences in effect size between the two methodologies. For example, in ultrasound, averaged differences in tongue position between the two harmonic contexts reach up to 2.5 mm. In EMMA, the maximal average difference was 1.40 mm. Such differences in magnitude are at least in part due to the fact that ultrasound allows access to almost the entire surface of the tongue as opposed to the position of a few flesh-points with EMMA. Details of the ultrasound data and its relation to the EMMA data will not be discussed due to space limitations (see Benus 2005).

5.3 Synthesis: dynamical links between continuity and discreteness

The main generalization from our study is that continuous differences in tongue body posture of stem-final vowels are linked to alternations in discrete suffix form. A more advanced transparent vowel selects a front suffix and a less advanced one selects a back suffix. We now propose a model which derives this generalization. Our presentation will proceed in two steps. To capture differences in degree of tongue position for stem-final vowels, we first model harmony between stem vowels as perturbations of consecutive vowel gestures due to coarticulation. These perturbations effect differences in the spatial targets of stem-final vowels in different environments. We then model the relation between these continuous differences and the binary ([±back]) nature of the suffix alternations.

In the proposed model of stem-internal harmony, articulatory gestures are formalized using point-attractor dynamics. As discussed earlier, a gesture has both spatial and temporal dimensions. Since in vowel harmony the combinatorial patterns on vowels are described in terms of the spatial dimension of constriction location (CL), we will focus on CL here. For any given gesture, we model the dynamics of CL using a monostable potential of the form $V(x) = \alpha(x - x_0)^2$, where α expresses the strength with which a given gesture imposes its control over the tongue body articulator and x_0 represents the target CL value. For exposition, we choose $CL = 2$ as a representative value for front or [–back] vowels and $CL = -2$, a constriction farther inside the vocal tract, as a representative value for [+back] vowels. We can visualize the movement of the tongue body articulator toward its target location by a particle left in the potentials of Fig. 10.

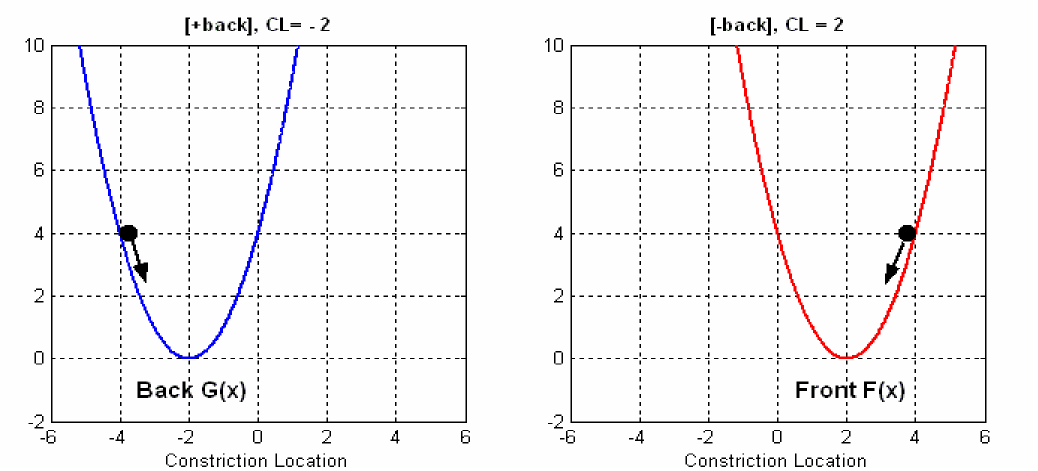
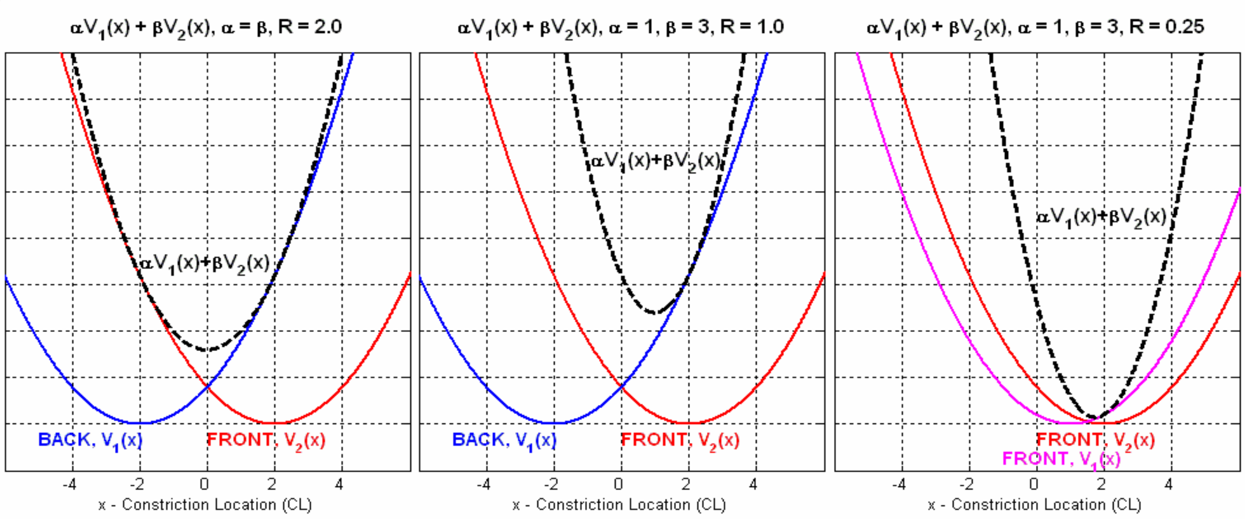


Fig. 10 Dynamics for constriction location.

A universal characteristic of speech is that the gestures of consonants and vowels overlap in time. This is known as coarticulation or coproduction. As a result of this overlap, gestures sharing an articulator exert spatial influences on each other. For example, when a back vowel is followed by a front vowel, /a-i/, the tongue body posture of the front vowel is relatively retracted due to the demands on the tongue body from the overlapping back vowel. Such spatial influences among gestures are known as blending. As first shown by Öhman’s (1966), vowels blend even when they are separated by consonants (see also Fowler 1983, Manuel & Krakow 1984, Saltzman & Munhall 1989). This blending of vowel gestures in consecutive syllables is a core component of our vowel harmony model. Consider two vowel gestures in adjacent syllables as in the stem *papír*. The first is a back vowel with the potential $V_1(x) = \alpha(x+2)^2$, and the second is a front vowel with the potential $V_2(x) = \beta(x-2)^2$. We focus here on the effects of blending on the second or stem-final vowel.[5] The simplest hypothesis for formalizing the perturbation of the second gesture due to the presence of the first vowel is to take the linear combination of the unperturbed potentials where both gestures contribute equally to the blended output for the second gesture ($\alpha = \beta$). The potential in the left panel of Fig. 11, marked with the dashed line, represents the perturbed front vowel gesture under this blending hypothesis.


$\alpha V_1(x) + \beta V_2(x)$, $\alpha = \beta$, R = 2.0

$\alpha V_1(x) + \beta V_2(x)$, $\alpha = 1$, $\beta = 3$, R = 1.0

$\alpha V_1(x) + \beta V_2(x)$, $\alpha = 1$, $\beta = 3$, R = 0.25

Fig. 11 Blending of two gestures represented with potentials $V_1(x)$ and $V_2(x)$.

---

[5] See Benus (2005) on how this model can also capture bi-directional blending in V-V sequences.

It can be seen that if $\alpha = \beta$, the attractor of the resulting potential would be at the midpoint between the FRONT and the BACK attractors. However, for stems like *papír*, our experiments show that the second vowel is a slightly retracted version of a front vowel rather than a vowel with a tongue body posture midway between a front and back vowel. Therefore, a minimal extension of the blending function is required so that the two potentials influencing the blended output of the second gesture carry different weights. This is a reasonable extension because our aim is to model the perturbation of the second gesture due to the first gesture. As the second gesture is produced, the strength of the control imposed on the tongue by the first gesture fades away while that of the second gesture increases. A simplified but adequate way of capturing this in our blending model is to increase the weight of the second gesture relative to that of the first (see also Fowler & Saltzman 1993 on a related notion of gestural strength). This is shown in the middle panel of Fig. 11 where now the gesture of the second front vowel is weighted more than that of the first, back vowel ($\alpha = 1$, $\beta = 3$). Consequently, the result of blending, the potential $\alpha V_1(x) + \beta V_2(x)$ shown with the dashed line, has its minimum tilted more toward the attractor of the second, front gesture. The right panel of Fig. 11 shows the result of blending on the second or stem-final vowel for a stem with two front vowels such as *emir* 'emir'. In this example, the first vowel is slightly less advanced ($CL = 1$) than the second one ($CL = 2$). Given that both V_1 and V_2 are front vowels, the difference between their respective attractors is small and the displacement of V_2 from its canonical horizontal position is minimal.

Gestural blending effects perturbations in vocalic tongue body position of stem-final vowels, e.g. the /í/ in *papír* is retracted to some degree because it follows a back vowel. The degree to which a vowel is retracted is captured by the parameter *R*, the difference between the attractor position of the vowel gesture after blending and its attractor position before blending.

Our experiments indicate that fine differences in articulatory retraction of transparent vowels are linked to the discrete alternation in suffix form. A more advanced transparent vowel selects a front suffix and a less advanced one selects a back suffix. How can small differences in articulation be related to a categorical ([±back]) alternation in suffixes? Informally, the relation obtained between degree of advancement of the stem-final vowel and suffix selection is nonlinear. Small changes in the former can cause large (nonlinear) changes in the latter. We now formalize this idea using nonlinear dynamics.

A first step in a dynamical model of a natural system is mapping the macroscopic observables to attractors of a hypothesized model underlying that system. In the case of vowel harmony, the relevant macroscopic observable is that suffix vowels alternate between a front and a back version. In the proposed model for suffix selection, the two discrete forms of an alternating suffix (e.g. Dative -*nak* vs. -*nek*) should correspond to the attractors of a dynamical system. In order to model the dependence between the continuous parameter of retraction degree R of the stem-final vowel and the discrete form of the suffix, we require that the choice of the suffix attractor be modulated by variation in the control parameter R. Following the discussion in section 2, these ideas can be stated in the form of equation $\dot{x} = f(x, R) + Noise$.

Our goal now is to determine a good candidate for the function $f(x,R)$. A proposed dynamical model of some phenomenon is a good model to the extent that aspects of the phenomenon in question correspond well with qualitative properties of its mathematical formulation. An appropriate dynamical system for the suffix alternation is required to have a bistable potential to capture the presence of two stable forms of a suffix, front and back. A polynomial of degree less than three allows for at most one attractor (Arnold 2000). Hence, the simplest model for suffix choice can be specified by a cubic polynomial. A good candidate for $f(x,R)$ is the 'tilted' anharmonic oscillator whose dynamics are described by

$f(x,R) = \lambda R + x - x^3$ where λ is a factor linearly proportional to *R* (Gafos, in press). Since $\dot{x} = f(x,R) = -dV(x)/dx$, we can compute the potential landscape $V(x) = -\lambda R - x^2/2 + x^4/4$ by integrating *f(x,R)*. Using concepts from section 2, the value of the constriction location for a suffix vowel is interpreted by the position of a particle running downhill in this potential and the asymptotic behavior of x in this equation can be visualized by looking at the simulations shown in Fig. 12. In these plots, the control parameter *R* varies between 0.2 and 1.2, corresponding to minimal and maximal retraction, respectively.

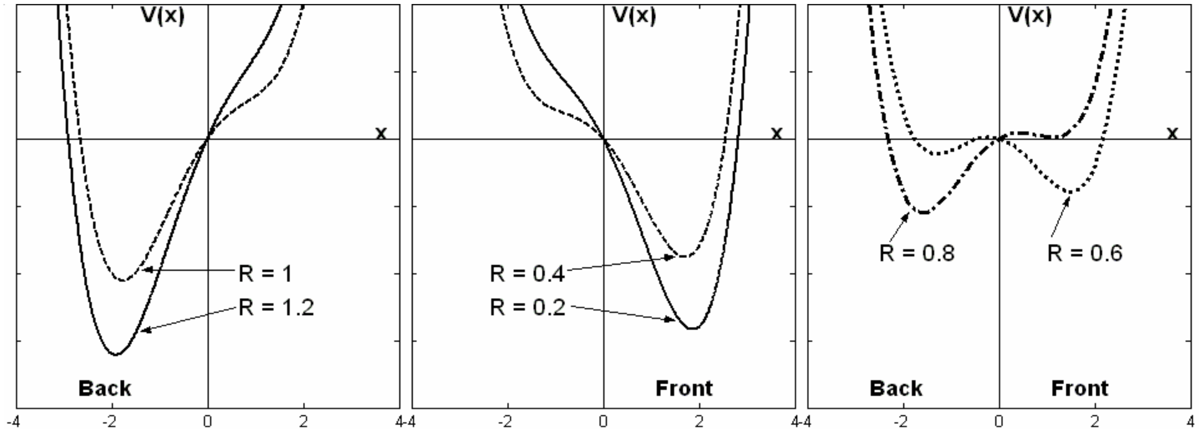


Fig. 12 Suffix form as a function of retraction degree *R*.

The graph on the left simulates suffix selection in stems like *Tomi*, which take back suffixes. We observed experimentally that in such stems the transparent vowel is retracted. In our model, retraction enters the dynamics via *R*. The function *f(x,R)* for a range of *R* values, $R \approx 1$ (significant retraction), provides a potential *V(x)* with an attractor close to the value of $CL = -2$ (BACK), corresponding to the back variant of the suffix. The probability that a particle left in this potential ends up in the vicinity of the BACK attractor is very high. Because the position of the particle represents the [±back] form of the suffix, it is predicted that the suffix is back, e.g. *Tomi-hoz*.

The graph in the middle panel shows how the potential *V(x)* changes for stems whose final vowels show minimal or no retraction like *emir*. For minimal retraction, modeled as $R \approx 0.2$, a qualitative change is evident in the shape of *V(x)*. The BACK attractor has been replaced by a FRONT attractor. A stem with minimal retraction of its final vowel is thus predicted to select front suffixes, e.g. *emir-hez*.

The right panel in Fig. 12 shows the behavior of the system for intermediate values of the control parameter *R* ($R \approx 0.7$). In nonlinear dynamics, a change from one macroscopic state of the system to another implies an intermediate stage of fluctuation. We see that there are now two minima representing the presence of two stable states, FRONT and BACK. For intermediate *R* values thus our model predicts that the suffix can vary between a front and a back version. To see this, we must consider the effects of noise and initial position of the particle. For example, consider a particle at a position around (0,0) in any of the potentials of the right panel. Due to the random kicks introduced by the fluctuations, the particle ends up in either the FRONT or the BACK attractor. We illustrated this with analytical and simulation methods described in section 2.

Before turning to the specifics of this prediction, consider what Fig. 12 tells us about the relation between the control parameter *R* and the order parameter of suffix quality. Equal changes in the control parameter do not always effect comparable changes in the order parameter. For example, both $R = 1.2$ and $R = 1$ result in qualitatively the same potential with the single BACK attractor, albeit with different stability. But as *R* changes from $R = 1$ to $R = 0.8$ or from $R = 0.4$ to $R = 0.6$ the potential changes qualitatively from a monostable regime to a bistable regime. Hence, a change of R by $\Delta = 0.2$ leaves the qualitative form of the system unaltered within a certain region of the control parameter values. But within a different region of control parameter values, a change of the same magnitude causes a qualitative change in the behavior of the system. This is another illustration of the fundamental property of nonlinearity.

Our model then predicts vacillation in suffix selection for intermediate retraction degrees. As discussed earlier, we do in fact find two sources of vacillation in Hungarian: Be stems where a back vowel is followed by the low /e/ (e.g. *hotel*) and BTT stems where a back vowel is followed by two transparent vowels (e.g. *aszpirin*, *mamicsi*). For Be stems, Benus (2005) argues that the low and relatively less front tongue body posture for /e/ allows for only limited retraction when /e/ blends with a preceding back vowel. Effectively, /e/ can be retracted less than the other transparent vowels /i í é/. This provides a basis for capturing the difference between Be stems and stems where the back vowel is followed by some other transparent vowel.

We concentrate here on the explanation that our model provides for the robust generalization that BTT stems such as *mamicsi* are more likely to select front suffixes than BT stems such as *mami*. In our model, greater tongue body advancement (i.e. smaller *R* values) of the stem-final vowel corresponds to greater probability of selecting front suffixes. If the final vowel in stems *mamicsi* can be shown to be less retracted compared to the final vowel in stems like *Tomi*, this prediction of the dynamic model would be confirmed.

Fig. 13 shows how our model derives the retraction degree of the final vowel in BTT stems such as *mamicsi*. Since in /a-i-i/ there are two pairs of adjacent vowels, our figure shows two panels corresponding to the two blending sites. On the left panel, we show the blending between /a/ and the following /i/, where /i/'s potential after blending is shown with the dashed line. We see that the blended gesture is retracted compared to the canonical /i/ gesture. Its attractor is at CL = 1, with a retraction degree $R_i = 1$.

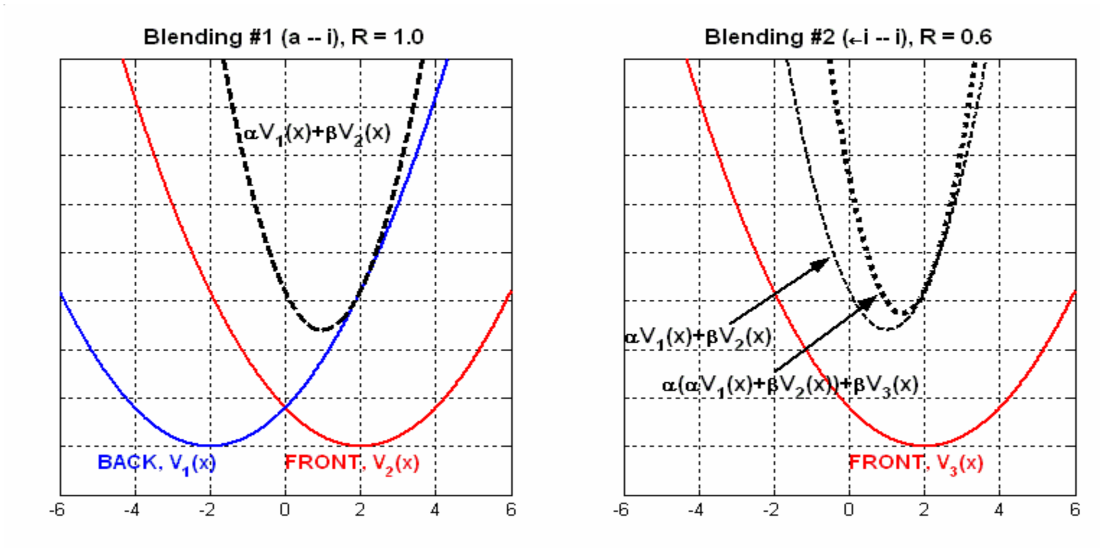


Fig. 13 Stem-internal blending in BTT stems.

If the perturbed /i/ vowel from the first blending was the stem-final vowel, then for its value of R our model for suffix selection would yield a suffix with a back vowel. As a result, a stem like *mami* is predicted to select back suffixes, consistent with the data. However, the /i/ in the second syllable of *mamicsi* is not stem-final because it is followed by another vowel. The right panel of Fig. 13 shows the second blending between the first retracted /i/, the output of the first blending, and the final /i/ vowel. This blending yields the potential shown with the dotted line whose attractor is at $CL = 1.4$. The degree of retraction of the stem-final vowel in *mamicsi* is thus $R = 0.6$. The main point is that, effectively, the second transparent vowel is less retracted than the first ($R_{T1} = 1.0$, $R_{T2} = 0.6$). As shown in Fig. 12 the potential *V(x)* for the suffix vowel following a stem-final vowel with $R = 0.6$ is bistable, but with a bias toward the FRONT attractor. The probability density function corresponding to this potential predicts that the probability of x being in a region around the FRONT attractor is higher than the probability of x being around the BACK attractor. This was illustrated with simulations in section 2, Fig. 4. Hence, the difference in the potentials for $R = 0.6$ and $R = 1$ translates into a difference in suffix selection between BTT and BT stems. The former are more likely to select front suffixes than the latter.

We thus see that by linking suffix selection with sub-categorical features of transparent vowels, the proposed model derives the difference between BT and BTT stems. All stem vowels participate in harmony since all vowel gestures undergo blending. Because there are more front vowels in BTT than in BT stems, the stem-final vowel in BTT stems is less retracted than in BT stems. In the proposed nonlinear model, this difference in retraction corresponds to qualitatively different suffix choices. We plan to test the empirically predicted values of retraction in BTT versus BT stems in a future study. These differences will in turn allow us to fine tune quantitatively the values of R generating a monostable versus a bistable potential (see Fig. 12).

To sum up, the starting point of this section was that small changes in tongue body constriction location of Hungarian transparent vowels are related to qualitative suffix alternations. This is the property of nonlinearity, a hallmark of complexity in natural systems in general and spoken language in particular. We presented a model that allows one to relate continuous phonetic distinctions to discrete phonological form using the mathematics of nonlinear dynamics. Our model accounts for the patterns of suffix selection in stems like *Tomi* (back suffix) vs. *Imi* (front suffix), and makes plausible predictions for the patterns vacillation (*hotel-nak/nek, mamicsi-nak/nek*). Overall, the proposed model provides an explicit link between the quantitative and qualitative aspects of the relevant patterns and makes specific predictions leading to new experimental studies.

6 Conclusion

At a broad level, the main proposal of this paper is that the formal language of nonlinear dynamics makes it possible to model the relation between the discreteness of phonological form and the continuity of phonetic substance in which that form is embedded. More specifically, the relation between discreteness and continuity, a fundamental conceptual problem in spoken language and other areas of cognitive science, is formalized by models relating two kinds of parameters – order parameters, describing the qualitative form of the system, and control parameters, which cause gradient drifts around qualitative states or nonlinear jumps from one qualitative state to another.

Acknowledgments

Thanks to… [to be added]. Both authors gratefully acknowledge support by NIH Grant HD-01994 to Haskins Laboratories.

References

Anderson, S. R. (1974). *The organization of phonology*. New York/London: Academic Press.

Arnold, V. I. (2000). Nombres d'Euler, de Bernoulli et de Springer pour les groupes de Coxeter et les espaces de morsification: le calcul de serpents. In É. Charpentier & N. Nikolski (Eds.), *Leçons de Mathématiques d'Aujourd'Hui* (pp. 61-98). Paris: Cassini.

Beckman, M. E. & Kingston, J. (1990). Introduction. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech* (pp. 1-16). Cambridge: Cambridge University Press.

Beer, R. D. (1995). Computational and dynamical languages for autonomous agents. In R. F. Port & T. van Gelder (Eds.), *Mind as Motion* (pp. 121-148). Cambridge, MA: MIT Press.

Benus, S. (2005). *Dynamics and transparency in vowel harmony*. Unpublished doctoral dissertation, New York University, New York.

Benus, S., Gafos, A., & Goldstein, L. (2004). *Phonetics and phonology of transparent vowels in Hungarian*. In P.M. Nowak, C. Yoquelet & D. Mortensen (Eds.), *Proceedings of the 29th Annual meeting of the Berkeley Linguistic Society* (pp. 485-497). Berkeley Linguistic Society.

Blumstein, S.E. (1991). The relation between phonetics and phonology. *Phonetica*, 48, 108-119.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart and Winston.

Browman, C. P., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3, 219-252.

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6, 201-251.

Browman, C. P., & Goldstein, L. (1990). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, 18, 299-320.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49, 155-180.

Browman, C. P. & Goldstein, L. (1995). Dynamics and articulatory phonology. In R. F. Port & T. van Gelder (Eds.), *Mind as Motion* (pp. 175-193). Cambridge, MA: MIT Press.

Charles-Luce, J. (1993). The effects of semantic context on voicing neutralization. *Phonetica*, 50, 28-43.

Charles-Luce, J. (1997). Cognitive factors involved in preserving a phonemic contrast. *Language and Speech*, 40, 229-248.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York: Harper and Row.

Clements, N. G. (1985). The geometry of phonological features. *Phonology Yearbook*, 2, 225-252.

Cohn, A. C. (1990). *Phonetic and phonological rules of nasalization*. Doctoral dissertation, University of California, Los Angeles. [Published as *UCLA Working Papers in Phonetics* 76].

Coleman, J. (1992). The phonetic interpretation of headed phonological structures containing overlapping constituents. *Phonology*, 9, 1-44.

Dinnsen, D. A. (1985). A re-examination of phonological neutralization. *Journal of Linguistics*, 21, 265-279.

Dinnsen, D. A. & Charles-Luce, J. (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics*, 12, 49-60.

Eco, U. (1984). *Semiotics and the philosophy of language*. Bloomington: Indiana University Press.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303-306.

Eubank, S., & Farmer, J. D. (1997). Probability, random processes, and the statistical description of dynamics. In L. Lam (Ed.), *Introduction to Nonlinear Physics* (pp. 106-151). New York: Springer-Verlag.

Farkas, D., & Beddor, P. (1987). Privative and equipollent backness in Hungarian. In A. Bosch, B. Need & E. Schiller (Eds.), *23rd Annual Regional Meeting of the Chicago Linguistics Society. Part Two: Parasession on autosegmental and metrical phonology* (pp. 90-105). Chicago: Chicago Linguistics Society.

Fodor, J. A., & Pylyshyn, Z. W. (1981). How direct is visual perception? Some reflections on Gibson's 'ecological approach'. *Cognition*, 9, 139-196.

Fougeron, C., & Steriade. D. (1997). Does deletion of French schwa lead to neutralization of lexical distinctions? In *Euro-Speech 1997: Proceedings of the 5th European conference on speech communication and technology, Vol. 7* (pp. 943-946). University of Patras, Greece.

Fowler, C. (1983). Converging sources of evidence for spoken and perceived rhythms of speech: cyclic production of vowels in sequences of monosyllabic feet. *Journal of Experimental Psychology*, 112, 384-412.

Fowler, C., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech*, 36, 171-195.

Freidlin, M., & Wentzell, A. (1984). *Random perturbations of dynamical systems*. New York: Springer-Verlag.

Gafos, A. (In press). Dynamics in grammar: Comments on Ladd an Ernestus & Baayen. In L. Goldstein, D. Whalen & C. Best (Eds.), *Varieties of phonological competence (Laboratory Phonology 8)*. Berlin, New York: Mouton de Gruyter.

Gafos, A. & Benus, S. (2003). *On neutral vowels in Hungarian*. In M.-J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences [15th ICPhS]* (pp. 77-80). Universitat Autònoma de Barcelona, Spain.

Goldsmith, J. (1976). An overview of autosegmental phonology. *Linguistic Analysis*, 2, 23-68.

Goodwin, B. C. (1970). Biological stability. In Waddington, C. H. (Ed.), *Towards a Theoretical Biology, vol. 3: Drafts* (pp. 1-17). Chicago: Aldine Publishing Company.

Haken, H. (1977). *Synergetics, An introduction*. Heidelberg: Springer-Verlag.

Haken, H. (1990). Synergetics as a tool for the conceptualization and mathematization of cognition and behavior – How far can we go?. In H. Haken & M. Stadler (Eds.), *Synergetics of Cognition* (pp. 2-31). Heidelberg: Springer-Verlag.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42, 335-346.

Harshman, R. A., Ladefoged, P., & Goldstein, L. (1977). Factor analysis of tongue shapes. *Journal of the Acoustical Society of America*, 62, 693-707.

Hayes, B. (2004). Stochastic phonological knowledge: the case of Hungarian vowel harmony. Talk presented at New York University, April 23, 2004.

Higham, J. D. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review* (Education Section), 43, 525-546.

Hulst, H.G. van der. (1988). The geometry of vocalic features. In H. van der Hulst & N. Smith (Eds.), *Features, segmental structure and harmony processes, Vol. 2* (pp. 77-125). Dordrecht: Foris.

Iskarous, K. (In press). Edge detection and shape measurement of the edge of the tongue. *Clinical Linguistics and Phonetics.*

Jong, J. K.de, Lim, B., & Nagao, K. (2001). Phase transitions in a repetitive speech task as gestural recomposition. *Journal of the Acoustical Society of America*, 110, 2657.

Jongman, A. (2004). Phonological and phonetic representations: The case of neutralization. In A. Agwuele, W. Warren & S.-H. Park (Eds.) *Proceedings of the 2003 Texas Linguistics Society Conference*, (pp. 9-16). Somerville, MA: Cascadilla Proceedings Project**.**

Kaun, A. (1995). *The typology of rounding harmony: An Optimality Theoretic approach*. Doctoral dissertation, University of California, Los Angeles. [Published as UCLA Dissertations in Linguistics, No. 8.].

Keating, P. A. (1988). Underspecification in phonetics. *Phonology*, 5, 275-292.

Keating, P. A. (1990). Phonetic representations in generative grammar. *Journal of Phonetics*, 18, 321-334.

Kelso, J.A.S. (1984). Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology*, 15, R1000-R1004.

Kelso, J.A.S., Saltzman E. L., & Tuller, B. (1986). The dynamical perspective on speech production: data and theory. *Journal of Phonetics*, 14, 29-59.

Kim, H., & Jongman, A. (1996). Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics*, 24, 295–312.

Kosslyn, S. M. (1978). Imagery and internal representation. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 217-257). Hillsdale, NJ: Erlbaum Associates.

Liberman, A. M., Harris, K.S, Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology,* 54, 358-368.

Manuel, S. Y., & Krakow, R. A. (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. In *Haskins Laboratories Status Report on Speech Research* 77/78, (pp. 69-78). Haskins Laboratories, New Haven, CT.

Ohala, J. (1994). Towards a universal, phonetically-based, theory of vowel harmony. In *Proceedings of International Conference on Spoken Language Processing* [September 18-22, 1994, Yokohama] (pp. 491-494).

Öhman, S. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, 39, 151-168.

Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I., & Jackson, I. (1992). Electromagnetic midsaggital articulometer (EMMA) systems for transducing speech articulatory movements. *Journal of the Acoustical Society of Americ*a, 92, 3078-3096.

Percival, I., & Richards, D. (1982). *Introduction to dynamics*. Cambridge: Cambridge University Press.

Piroth, H.G., & Janker, P.M. (2004). Speaker-dependent differences in voicing and devoicing of German obstruents. *Journal of Phonetics*, 32, 65–80.

Prince, A., & Smolensky, P. (1993). *Optimality Theory: Constraint interaction in generative grammar.* Ms., Rutgers University and University of Colorado. [Published by Blackwell, 2004].

Prince, A. & Smolensky, P. (1997) Optimality: from neural networks to universal grammar. *Science*, 275, 1604-1610.

Port, R. F., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, 17, 257-282.

Recasens, D. (1999). Lingual coarticulation. In W.J. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques in speech production* (pp. 78-104). Cambridge: Cambridge University Press.

Repp, B. H., & Liberman, A. H. (1987). Phonetic category boundaries are flexible. In S. Harnard (Ed.), *Categorical perception* (pp. 89-112). Cambridge: Cambridge University Press.

Ringen, C.O., & Kontra, M. (1989). Hungarian neutral vowels. *Lingua*, 78, 181-191.

Sagey, E. (1986). *The representation of features and relations in nonlinear phonology*. Doctoral dissertation, MIT, Cambridge, Massachusetts. [Published 1991, Garland, New York.]

Saltzman, E. (1995). Dynamics and coordinate systems in skilled sensorimotor activity. In: R. F. Port & T. van Gelder (Eds.), *Mind as Motion* (pp. 149-174). Cambridge, MA: MIT Press.

Saltzman, E., & Munhall, K. (1989). A dynamic approach to gestural patterning in speech production. *Ecological Psychology*, 1(4), 333-382.

Siptár, P., & Törkenczy, M. (2000). *The phonology of Hungarian*. Oxford: Oxford University Press.

Stetson, R. H. (1951). *Motor phonetics*. Amsterdam: North-Holland.

Stevens, K. N. (1972). The quantal nature of speech: evidence from articulatory-acoustic data. In E. David & P. Denes (Eds.), *Human communication: A unified view* (pp. 51-66). New York: McGraw-Hill.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics,* 17, 3-45.

Stone, M. (1997). Laboratory techniques for investigating speech articulation. In J. Hardcastle & J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 11-32). Oxford: Blackwell.

Turvey, M. T. (1990). Coordination. *American Psychologist*, 45, 938-95

Trubetzkoy, N. S. (1969). *Principles of phonology*. Berkeley & Los Angeles: University of California Press.

Vago, R. M. (1980). *The sound pattern of Hungarian*. Washington: Georgetown University Press.

Wood, S. (1979). A radiographic analysis of constriction location for vowels. *Journal of Phonetics*, 7, 25-43.