

Dynamics: the non-derivational alternative to modeling phonetics-phonology

Adamantios I. Gafos

January 30, 2003

[Paper presented at the Laboratory Phonology 8 meeting, New Haven, CT, 27 June 2002]

1. Introduction

The derivational view of phonetics-phonology (Ladd, this volume) expresses an intuition that seems valid, namely, that there is a distinction to be made between quantitative and qualitative aspects of phonetics-phonology. Incomplete neutralization (Ernestus and Baayen, this volume) and other phenomena like it indicate that the specific way of drawing that distinction is too rigid. At the same time, these phenomena underscore the need for a different formal language, where discrete and continuous aspects of phonetics-phonology can interact. A way of reconciling the core intuition of the derivational view with phenomena like incomplete neutralization is put forth by using the mathematics of nonlinear dynamics. This allows one to integrate the continuous and the discrete without the additional postulate that phonology be derivationally antecedent to phonetics.

2. Two views of phonetics-phonology

How are the qualitative aspects of phonological competence related to their variable and continuous phonetic manifestation? This question defines the so-called ‘phonetics-phonology problem’ and it has been one of the central themes of laboratory phonology (Beckman and 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 aspects of performance, as shown by parallel research in vision (Haken 1990), biological coordination (Turvey 1990), agent-environment interaction (Beer 1995) and other domains.

At a broad level, there are two views on the formalization of theories aiming to address this central question. One view, firmly established with the development of generative phonology (Chomsky and Halle 1968) and subsequently elaborated and refined in important ways, posits that the relation between qualitative and quantitative aspects of phonetics-phonology consists of a process of translation from discrete symbols to continuous physical properties of an articulatory and acoustic nature. In Ladd’s words, “we need to think of phonetic realization as a mapping between a categorical symbolic representation and a quantitative physical signal” (Ladd, this volume:6). This is the view in the background of most current work in phonetics-phonology and cognitive science in general, e.g., see the notion of transducer in Fodor and Pylyshyn (1981) and also Harnad (1990).

An alternative, relatively more recent and less widely explored view builds on the mathematics that can express both the discrete and the continuous aspects of complex systems, the so-called nonlinear dynamics (see Smolensky 1988, Port and van Gelder 1995 for a proposal and a sample of applications of the dynamical view in cognitive science, respectively). In phonetics-phonology, a precedent is Browman and Goldstein’s (1986 *et seq*) research programme. An important contribution emerging from Browman and Goldstein’s work is a formally explicit theory of dynamically defined phonological representations. Roughly speaking,

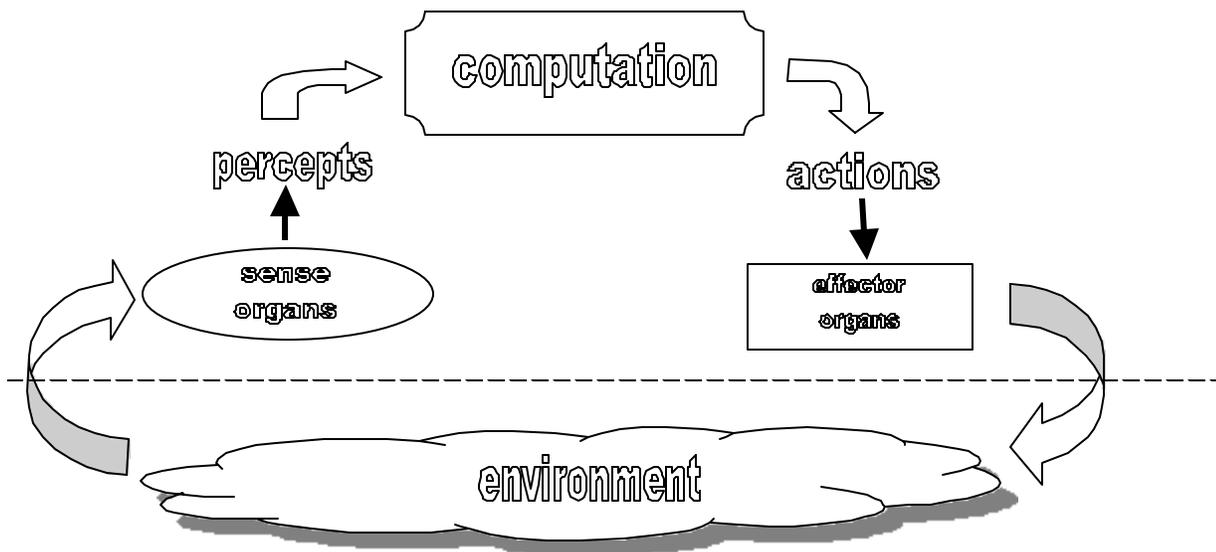
this theory implies that the atoms of phonological representations must be construed as unfolding in time (gestures) and that universal as well as language-particular principles may refer to this temporal dimension of phonological form (Browman and Goldstein 1986, 1995, Gafos 2002).

The goal in this paper is to broaden the argument for the dynamical view by focusing on a special case of the fundamental question here, the subtle context-dependency of phonological neutralization. As I discuss, in certain well-documented cases of phonological neutralization grammatical requirements interact with variable environmental conditions (here, speakers intentions to convey contrasts). This turns out to be a problem for the derivational view of phonetics-phonology. The specific aim is to show that a dynamics model predicts this context dependency of neutralization, an aspect of the problem that has remained outside the scope of previous models.

3. The problem: final devoicing

To state the problem in most general terms, it is useful to review the three main components of cognition: perception-computation-action, as shown in (1) (Cariani 1989). For example, in a task where a listener is asked to produce the plural of a spoken word, the perceptual system identifies the singular form, say, the percept [glik], the grammar computes the plural form [gliks], and finally the output computed by the grammar is implemented as vocal-tract action.

(1) Main components of biological cognitive agents: perception, computation, production



A fundamental fact left out from this description of the perception-computation-action loop is that the cognitive system is embedded in a continuously varying environment. Moreover, all three components of the system have the remarkable capacity to deal with various sources of variability in that environment.

Consider two prototypical examples from production and perception. It is well known that the timing characteristics subserving various segmental contrasts are dependent on speech rate. For example, Summerfield (1981) shows that the VOT boundary (onset of voicing relative to oral release) between voiceless and voiced consonants changes as a function of speech rate. As rate increases, speakers' productions of voiceless and voiced consonants shift towards shorter

values of VOT. In turn, listeners are sensitive to such variations, and adapt flexibly to different rate conditions. Another example is illustrated with the durational boundary between single and double consonants in examples like “topic” and “top pick”. This boundary is not invariant but depends on the rate of the utterance these tokens are part of. The faster the rate, the lower the boundary value. Listeners are sensitive to this rate-dependent change in the signal. A given silence duration is judged differently depending on the rate of the utterance. Similar results hold for the distinction between /s/-/ss/ in Japanese, a language with distinctive consonant length (see Miller 1981 for a review).

So far then we see that production and perception are stable in that varying some external parameter leaves the qualitative nature of the system, the distinct categories, unaltered. These systems are also flexible, because they adapt to varying environmental requirements, such as speaking fast or slow.

Next, consider an example from the cross-linguistically common phenomenon of final devoicing. The phonological description of final devoicing or neutralization is simple. In certain languages, obstruents are voiceless syllable-finally (Bloomfield 1933:218, Trubetzkoy 1969:213). See (2) for representative examples from German and Ernestus and Baayen (this volume) for Dutch.

(2) <i>Rad</i>	‘wheel’	[ʁat]	(nom.)	[ʁadəs]	(gen.)	inferred underlying form = /ʁad/
<i>Rat</i>	‘advice’	[ʁat]	(nom.)	[ʁats]	(gen.)	inferred underlying form = /ʁat/
<i>Bund</i>	‘association’	[bunt]	(nom.)	[bundəs]	(gen.)	inferred underlying form = /bund/
<i>bunt</i>	‘colorful’	[bunt]	(sing.)	[buntə]	(pl.)	inferred underlying form = /bunt/

The situation is more complex in the phonetics of neutralization. There are two main results. First, neutralization is incomplete in that the [t] in [bunt] ‘association’ is not identical to the [t] in [bunt] ‘colorful’. Even though both forms are transcribed with a voiceless [t], the mean of the variable indexing lack of voicing differs between the underlying [-Voiced], [+Voiced] consonants. The latter’s mean is slightly shifted toward less extreme values of devoicing or toward more “slight voicing” in Ernestus and Baayen’s (this volume) terms. Specifically, differences can be observed in vowel duration, consonant closure duration, and period of glottal pulsing during the consonantal closure. See, among others, Dinnsen (1985) for a review of other instances of incomplete neutralization, Dinnsen and Charles-Luce (1984) on Catalan final devoicing, Fougeron and Steriade (1993) on French schwa elision, and Charles-Luce (1993, 1997) on Catalan voicing assimilation and English flapping, respectively.¹

Second, neutralization shows a subtle dependency on the communicative context. This can be illustrated with the following task, from Port and Crawford (1989). In one experimental condition, speakers are asked to read a list of words in isolation. In the another condition, speakers are asked to read sentences like *Ich habe Rat(Rad) gesagt; nicht Rad(Rat)* (“I said Rat(/Rad) not Rad(/Rad)”) while a German assistant, who is present in the experimental setting, is assigned the task of writing down the order of the test words in such sentences. In this second

¹ Fourakis and Iverson (1984) ascribe the incompleteness of neutralization to “hypercorrection under linguistically artificial conditions [AG: orthography in word list reading]” (149). But Catalan, a language where incomplete final devoicing has been documented, lacks an orthographic distinction between word-final underlying voiced and voiceless stops (see references in the text). See also Charles-Luce (1985:318-9), Port and Crawford (1989:258-9), and Ernestus and Baayen (this volume) for related discussion.

condition, then, speakers are encouraged by the context to convey the contrast more than in the word list reading condition. The observed result is a stronger version of incomplete neutralization than in the word list reading condition (where no assistant is present). This is to say that the means of the variables indexing voicing shift even more toward less extreme values of devoicing (Port and Crawford 1989, see also Charles-Luce 1985).

The incompleteness of final devoicing and its systematic dependence on context are characteristic of the flexibility and stability of the phonetics-phonology system. On the one hand, there is a consistently reproducible aspect of the phonetics-phonology of German, identified with final devoicing (stability). On the other hand, the phonetics-phonology system is flexible in allowing speaker's intentions to shift the phonetic output in ways that deviate slightly from the ideal grammatical optimum (flexibility).

Consider how the derivational view of phonetics-phonology deals with stability and flexibility, in general. The symbolic constructs of phonology are by definition stable – they are mental realities abstracted from the environment (axiomatic stability). The grammar is stable because its essential constructs are symbolic in nature. Flexibility enters the life of the phonetics-phonology system in phonetic implementation, after the grammar has computed an output or 'interface representation' in Ladd's terms. In phonetic implementation, symbolic units are translated to vocal tract action under different conditions – different speech rates, styles, social contexts, etc. – and environmental variables begin to introduce their effects.

However, the incompleteness of neutralization does not fit comfortably in this view. This is illustrated in (3). Final Devoicing changes the voicing value of the final obstruent in /bund/ to [-Voiced]. This eliminates the contrast between the final consonants of /bund/, /bunt/ at the output of phonology, exactly as a 'neutralization' rule should do. Consequently, phonetic implementation, whose role is to flesh out phonology's output as vocal-tract action, is now unable to deliver the differences observed in the realizations of /bunt/ versus /bund/.

(3) Rule of Final Devoicing, FD: [+Voiced, –Sonorant] → [-Voiced] / ___^σ

Underlying forms	Output of phonology		Vocal-tract action
/bunt/	[bunt]	→	[t], wide glottal aperture
/bund/	[bunt] (via FD)	→ ?	[t ⁺], less wide glottal aperture

It is clear that incomplete neutralization requires some revision of the standard phonology-phonetics view. Accordingly, the incompleteness of final devoicing has led to arguments for relaxing one of the foundational assumptions of that view, the ordering of phonology before phonetic implementation. See Dinnsen and Charles-Luce (1984:58), Slowiaczek and Dinnsen (1985:338) for Catalan and Polish, respectively.

Another approach is to apply Final Devoicing at the same time as phonetic implementation (Ernestus and Baayen, this volume; Port and O'Dell 1985). As Ernestus and Baayen observe, the main problem with this proposal is that phonology becomes indistinguishable from phonetic implementation. Final devoicing is an aspect of German phonology. By moving it to phonetic implementation, final devoicing must be reformulated in a different formal language, the language of phonetic implementation, using continuous mathematics. The proposal to be fleshed out here begins with the challenge of maintaining the distinction between qualitative versus quantitative aspects of phonetics-phonology by proposing an appropriate parameterization of the phonetics-phonology system.

The second, equally important characteristic of incomplete neutralization is its systematic dependence on the communicative context. To date, I am not aware of any previous formal treatment of this effect. This phenomenon is an example of what Liberman (1983) refers to as phonological systematicities which are “modulated by ... paralinguistic parameters” and which are “not well modeled as feature- or structure-changing rules” (271). The grammar output is quantitatively shifted by intentions, but intentions are not the kinds of primitives that are described as being part of the grammar – they are paralinguistic.

The challenge for the derivational view of phonetics-phonology is that, on the one hand, placing final devoicing in the phonology captures the fact that final devoicing is a qualitative property of German, but it cannot account for the flexibility of the phonetics-phonology system. On the other hand, moving final devoicing to phonetic implementation would allow it to be modulated by extra-grammatical, continuous factors but loses sight of the fact that final devoicing is an aspect of German phonology.

The alternative to be proposed here is a non-derivational (parallel) way of relating discrete aspects of the grammar and continuous, environmental variables. This promises to bypass the ordering problem, under the assumption that there is a coherent way to make continuity and discreteness coexist within the same formal language, and also that there is a way to at least describe and at best derive phenomena like incomplete neutralization. The mathematics of nonlinear dynamics satisfies the first assumption, as discussed in the next section. Subsequent sections take up the issue of deriving incomplete neutralization, using basic concepts of nonlinear dynamics.

Before leaving this section, I consider whether phonological models dealing with variability can be of help with the problem faced here. In a rule-based model (Chomsky and Halle 1968), we may consider ‘variable rules’ as in Sankoff (1987) or Cedergren and Sankoff (1974). In Optimality Theory (Prince and Smolensky 1993), we may consider the ‘stochastic evaluation’ method for constraint interaction as proposed in Boersma and Hayes (2000). To illustrate, in the latter model, if two constraints are sufficiently close on a rank scale, a small shift in their rank values can result in $C1 \gg C2$ or $C2 \gg C1$. In the specific example, the constraints are $C1 = \text{NOVOICEDCODA}$, $C2 = \text{FAITH(Voice)}$. Their variable ranking would give rise to underlyingly voiced obstruents being produced sometimes voiceless (when $C1 \gg C2$) and sometimes voiced (when $C2 \gg C1$).

These models deal with a different type of variation from that addressed in this paper. They deal with variation among *discrete* alternatives. In the present case, however, it is not that the voiced obstruent is produced sometimes voiced and sometimes voiceless. Rather, the mean value of voicelessness gradually drifts toward less extreme values, and it does so lawfully as a function of the communicative context. Hence, those models are inapplicable to this type of variation, which I will call lawful continuous variation.

There is, however, another class of models with the capacity of handling continuous dimensions, the so-called exemplar models of memory and categorization (Hintzman 1986). Recently, Pierrehumbert (2001, 2002) has developed an application of the exemplar paradigm to phonetics-phonology, with attention to variation and fine phonetic details in the realization of phonological categories. Specifically, in that application, variation in production is achieved by averaging and/or randomization over a set of memorized exemplars of a category, generating a so-called ‘echo’ of the category. The crucial observation here is that the variation involved in final devoicing has a systematic component, as changes in environmental variables result in

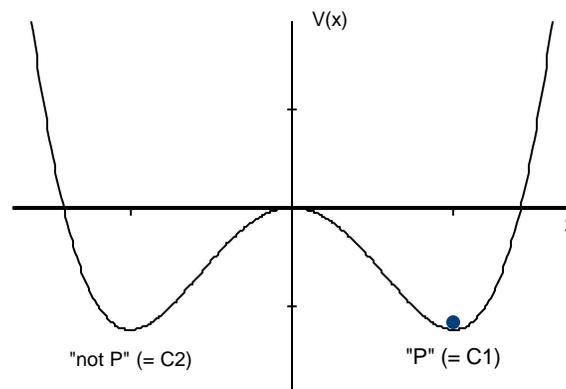
systematic gradual drifts toward more or less voicing. This context dependency is not accounted for by an averaging and/or randomization method, as in fact noted in Pierrehumbert (2002:19).

4. Phonetics-phonology in a dynamical setting

To develop a parallel view of phonetics-phonology, the essential insights of the field must be recast using the mathematics of nonlinear dynamics. Thus, phonetic categories, representations, constraints, and grammars must be given a dynamical formulation. For phonetic categories and representations, some of the foundational work in these domains has been couched in terms that are at least consistent with the dynamical approach. See Stevens (1972, 1989), Petitot-Cocorda (1985), Kingston and Diehl (1994), and references in section 2 on Browman and Goldstein's work. In this section, I focus on constraints and grammars. To anticipate, my specific proposal is that constraints are *attractors* and that grammars are *attractor landscapes*. Both notions are basic to nonlinear dynamics.

To begin, phonological constraints are formulated as competing *attractors* (Thompson and Stewart 2002:45). Attractors define preferred modes for the macroscopic parameters of phonology. For example, the constraints “BE VOICELESS” or “BE CORONAL” state preferred values for the phonological parameters of Voicing and Place of articulation. In (4), two competing constraints C1, C2 are depicted as two attractors; attractor 1: ‘have property P’, attractor 2: ‘not P’. Taking Voicing as an example and letting x represent the parameter of glottal aperture, the system can be in two states. Either it is “Voiceless”, it has property P, or it is “Voiced”, it does not have property P. The first state is represented with the minimum at some positive value of the glottal aperture variable x , and the second state with the minimum at the some negative value of the same variable (the actual values and their signs are not crucial in the present context).

(4) Phonological constraints as competing *attractors*



The figure in (4) represents the assumption that, in a language with a Voiceless/Voiced contrast, the Voicing parameter draws values from two recognizably distinct parts of its state space (the state space is the entire x -axis). It thus describes qualitatively distinct modes of the voicing system, or in other words it describes a dimension of macroscopic order in phonological form. For this reason, it is called an *order parameter* (Haken 1977).

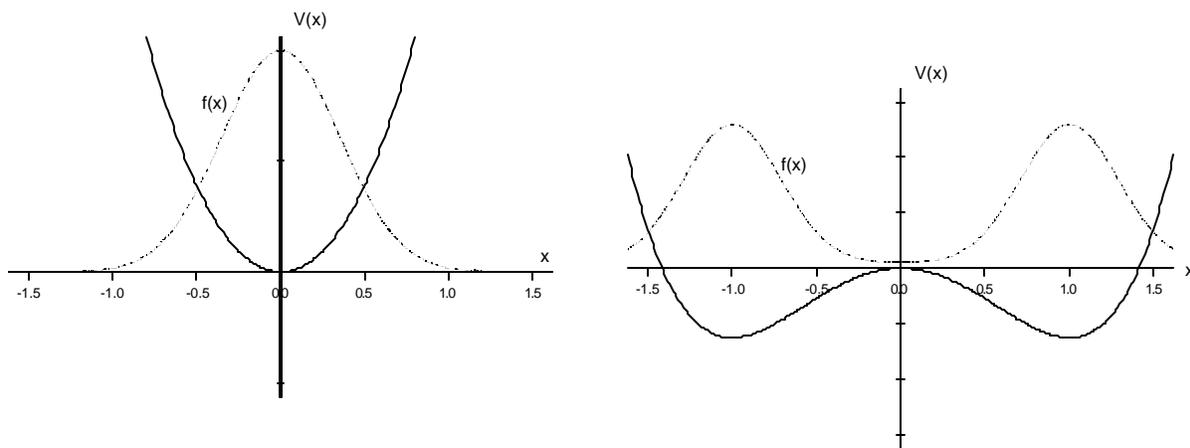
Intuitively, we may interpret the behavior of an order parameter by means of a ball moving in the potential $V(x)$ shown above. Clearly, the ball ends up in one of the two attractors,

the macroscopic observables of the system. The attractor landscape shown there is known as the ‘anharmonic oscillator’ and it is described by the potential function $V(x) = (-1/2)*(x^2) + (1/4)*(x^4)$.

Given that macroscopic order is expressed via order parameters and constraints referring to these, what is the relation between these parameters and traditional symbols? Specifically, what is the symbol [“Voiced] in the dynamical formulation of the voicing distinction? In the dynamical formulation, the symbol is inseparably linked with its phonetic substance. It is not derivationally antecedent to that substance and therefore it does 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). Next, how is the stability of macroscopic order achieved in a dynamical formulation of phonetics-phonology? Attractive modes are *dynamically stable*, that is, they exhibit small fluctuations around their mean states (the two minima shown above). Fluctuations are inevitable due to noise. Noise is inevitable because complex systems described by low-dimensional dynamics are coupled to various subsystems at a more microscopic level. In our case, the control of voicing, the microscopic level corresponds to the neuronal, aerodynamic and myodynamic subsystems (Titze 1988).

Following Haken (1977), I describe noise as a small, random perturbation force pushing the representative point of the system x , the position of the ball, back and forth randomly. Randomness introduces stochasticity and consequently we can only compute the *probability* for finding x within a given interval of values of x . This probability is described by the probability distribution function $f(x)$ multiplied by the length of the interval. Two probability distribution functions corresponding to two different potentials are shown in (5).

(5) $V(x)$ and probability distribution function $f(x)$ for two potentials

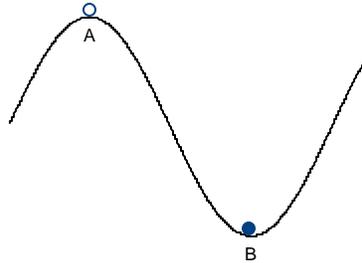


It can be seen that the probability to find the system around the mean state(s) of the attractor(s) is quite high. The probability to find the system at some other point decreases quickly as we move away from the mean states but it may not be zero. In short, the preferred modes of order parameters, the attractors, are resistant to noise in a probabilistic sense.

Noise is inherent to the process of modeling a phenomenon in dynamical terms and it can be used to generate predictions. Specifically, noise has a differential effect on the order parameters depending on the strength of the attractor. To illustrate, imagine the ball in the well of

a strong attractor. As a classic example, consider the ball at point B below. Here, noise has a small effect in causing minute perturbations around the mean state.

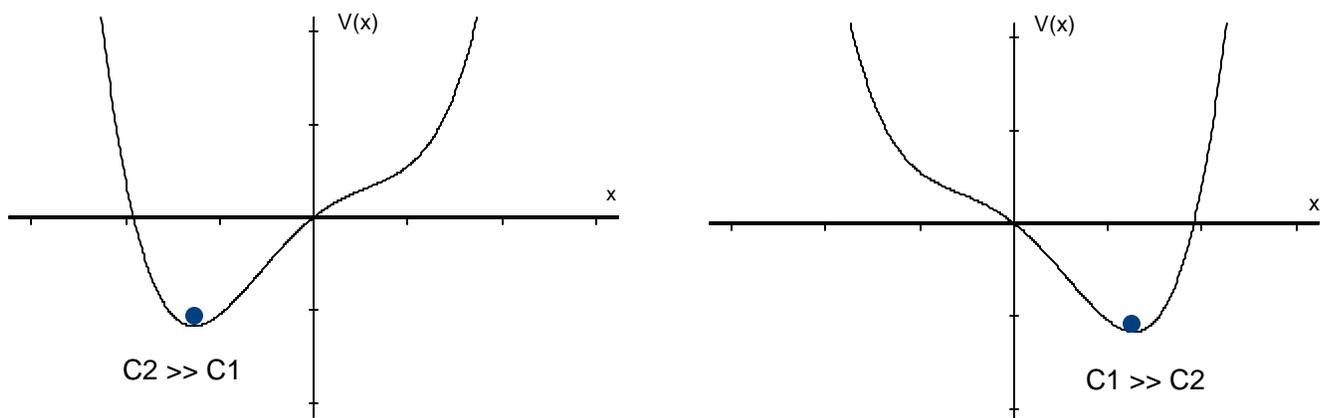
(6) Unstable (A) and stable (B) equilibria



Now, imagine what happens when the ball is put at point A. Due to random fluctuations, the ball ends up at the left or the right side. A is an unstable point. This illustrates that fluctuations can have dramatic effects at highly unstable regions. In dynamics, then, it is possible to exploit noise to discover the stable attractors of the system. Noise can be measured by the variance or standard deviation of some essential variable x around the attractive state. The more stable the attractor the smaller the deviation from the attractive state.

I now turn to a fundamental insight on grammars, namely, the idea that the qualitative aspects of linguistic form are the result of constraint optimization, and specifically the notion of constraint ranking. Both of these derive from Optimality Theory (Prince and Smolensky 1993). In the proposed model, constraint ranking is modeled as reorganization of the attractor landscape. This is illustrated in the figures below, which show two qualitatively different reorganizations of the attractor landscape in (4).

(7) Constraint ranking as reorganization of the attractor landscape – compare with 4



To do this, we adjust the so-called ‘control parameter’ k in the potential function $V(x) = k*x + (-1/2)*(x^2) + (1/4)*(x^4)$ which determines the tilt and direction of the potential (see

Tuller et al. 1994 for an application to perception). Thus in (7), the potential to the left corresponds to $k = 1$, the one to the right to $k = -1$, and the potential in (4) corresponds to $k = 0$ where the two constraints are unranked. It can be seen that the shift from $C1 \gg C2$ to $C2 \gg C1$ or vice versa implies an intermediate stage where $C1, C2$ are unranked (due to $-1 < 0 < 1$). Thus, grammar change implies variation. This is a corollary of the dynamical formulation of constraint ranking that seems consistent with the course of sound change (e.g., Lass 1997:64, 287, Sommerstein 1977:250-1). Moreover, it is possible to model fine, probabilistic variation in constraint ranking by smoothly varying the control parameter k . As k modifies the attractor landscape, the probability distribution function over that landscape changes accordingly, thereby modulating the probabilities of the different states the system may reside in. However, I cannot illustrate these consequences of the model in detail here.

In the remaining, I focus on two properties of nonlinear dynamics that are of critical importance in modeling complex systems in general and phonetics-phonology in particular. The first property is non-linearity. A system exhibits non-linearity when large or discontinuous changes can be observed in the behavior of that system as some control parameter varies smoothly. Examples in natural systems abound (Haken 1977, Winfree 1980). One such example from biological coordination is briefly mentioned here. Kelso (1984, 1995) 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 (this qualitative change is commonly referred to as a bifurcation by mathematicians, or a phase transition by physicists). In this example, then, gradual changes in cycling frequency drive the coordination system from one stable mode of coordination to another, anti-phase to in-phase. The phenomenon has been modeled in dynamical terms by Kelso and colleagues. For a recent review, see Wing and Beek (2002).

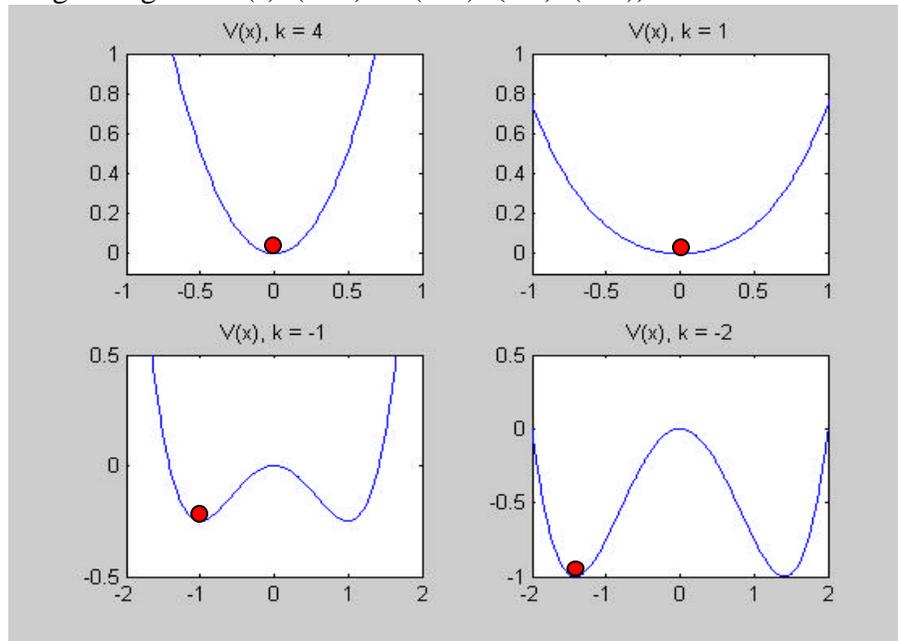
To return to phonetics-phonology, the formulation of constraint ranking given above exploits the property non-linearity. The systems in (4) and (7) are qualitatively different. They correspond to distinct Optimality Theoretic grammars, “ $C1, C2$ unranked” in (4) and “ $C1 \gg C2$ ”, “ $C2 \gg C1$ ” in (7). What makes this formulation of constraint ranking particularly relevant to phonetics-phonology is that it comes with a handle for driving the system from one qualitative state to another, as a consequence of varying the control parameter k . So from smooth, continuous variation in some control parameter, distinct grammars can emerge. In nonlinear dynamics, then, continuity and discreteness coexist and interact within a unified framework. By contrast, in a derivational phonetics-phonology, there is no way to express this interplay between continuity and discreteness. In such a model, variation in continuous or environmental parameters cannot affect the discrete aspects of phonetics-phonology. Phonologists working on the phonetic bases of phonological patterns have encountered (instances of) this limitation repeatedly. Steriade (1997) has expressed this most accurately and succinctly: “phonetic implementation has to live with prior decisions taken in the phonology” (1997:3). To generalize the same, in the derivational model, the continuous aspects of phonetics-phonology are enslaved by the discrete dimensions of the system. But as Browman and Goldstein have pointed out, there are clear cases of bi-directional interaction between the discrete and the continuous, or between the macro- and micro-levels of description in their terms (see Browman and Goldstein 1995).

Perhaps the most striking evidence that continuous or extra-grammatical parameters do affect the grammar derives from alternations sensitive to quasi-categorical or continuous

parameters. An example is vowel deletion under increase in the continuous parameter of rate, e.g. *parade*, *Toronto*, [p^hred], [t^hranto]. Crucially, it is not that the schwa is hidden in fast speech due to increased overlap with its surrounding consonantal gestures (Davidson 2001). Rather, the schwa is categorically deleted, and at least for some speakers this is a systematic result of the continuous scaling in the control parameter of rate. In the derivational view, since rate is an extra-grammatical parameter it cannot affect grammatical computation. In that view, rate can only effect gradient changes in the phonological output. This means that the vowel can shorten or lengthen (recall Ladd's 'continuous mathematics' for phonetic implementation), but not deleted categorically. Deletion is a symbolic operation that needs to be described in the language of 'discrete mathematics'. By contrast, in the dynamical view proposed here, it is possible to express this interplay between continuity and discreteness. In broad terms, the approach is the same with that developed later on with respect to incomplete neutralization. There are two qualitatively distinct modes, vowel-present and vowel-deleted. As the control parameter of rate is increased, the vowel-present mode becomes progressively less stable. When a critical value of rate is reached, a bifurcation in the system's dynamical behavior occurs and the output changes discontinuously to the vowel-deleted mode.

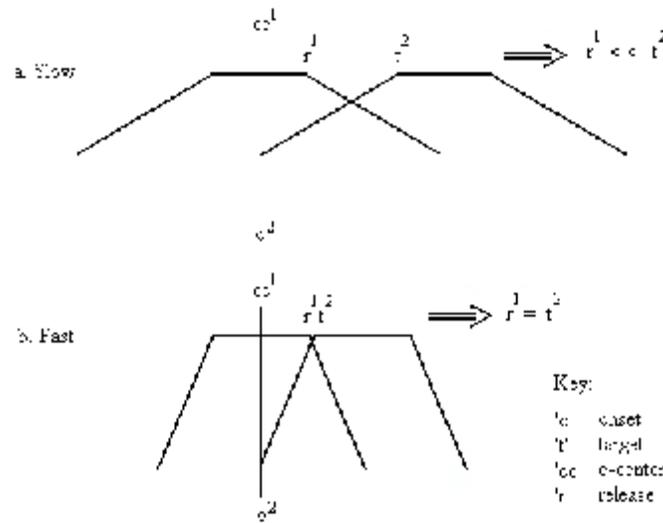
I now turn to the second property of nonlinear dynamics, scaling. Working on biological movement, Saltzman (1995) defines scaling as the "Lawful warping of a movement's form that can occur within parametric changes along performance dimensions such as motion rate and extent" (1995:152). To illustrate scaling, consider how the potential function in (8) changes as the control parameter k varies. However, focus now is not on the qualitative changes (the non-linearity property), but on the quantitative ones. Thus, for $k > 0$, see top row of the figure, as the control parameter changes the system retains a qualitative sameness of form – there is a single attractor. But variation in k does affect smoothly the attractor landscape. For example, from $k = 4$ to $k = 1$ there is a change in the strength of the attractor, as is evident from the flattening of the walls in the attractor's basin. Next, when k passes through a critical value (here, zero), suddenly a qualitative change occurs, and the system jumps to a bistable regime, with two attractors (bottom row). Here too, observe that as k changes from $k = -1$ to $k = -2$, the stable points (the two minima) drift apart smoothly. Overall, then, within the ranges of $k > 0$ or $k < 0$ the macroscopic form of the system remains unaltered, but variation in k does affect quantitative aspects of that form.

(8) Illustrating scaling with $V(x) = (-1/2)k(x^2) + (1/4)(x^4)$



An example of scaling from phonetics-phonology is illustrated in Gafos (2002). The relevant part of that work is concerned with modeling the effects of rate on final consonant clusters in Moroccan Arabic. Briefly, the relevant facts are as follows. At a normal speech rate, a final cluster of two heterorganic consonants is produced with an intervening acoustic release, known as an open transition, e.g. [kat^hb], active participle of ‘to write’. At faster rates, the intervening release is not present, [tb]. A computational simulation with a model of gestural dynamics shows that the temporal relation between the consonants that gives rise to the perceptual result of the open transition is the one in (9a). This relation is such that the *onset* of movement for the lips gesture for /b/ is initiated around the mid-point of the tip-blade gesture for /t/, the so-called *c-center* of /t/. As a consequence of this relation, the achievement of the *target* for the /b/ gesture, lip closure, takes place after the *release* of the /t/ gesture. There is, thus, a period of no constriction in the transition between /t, b/ that is identified as an open transition.

(9) Effects of rate on gestural movement in a sequence of two consonant gestures



The effect of increasing rate on the gestural kinematics is examined by increasing the control parameter of gestural stiffness, in Browman and Goldstein's (1995) model. In agreement with the data, the simulation shows that as rate is increased, the transitional acoustic release between the two consonants disappears – the consonants are produced in close transition, as shown schematically in (9b). This change from open to close transition is an observable consequence of scaling. Qualitatively the same timing relation between two consonants can have different acoustic consequences – the presence or absence of the transitional acoustic release – as the control parameter of speech rate changes smoothly.

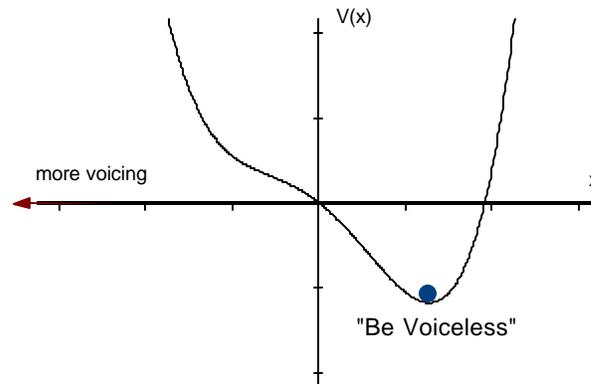
Next, I consider how the concepts introduced here can be applied to our specific problem, the incompleteness of neutralization and its dependence on the communicative context.

5. Grammar 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 (Kelso, Ding, and Schner 1992).

Consider the specific phenomenon addressed here, a language with final devoicing. The relevant macroscopic observable is the devoicing of final obstruents, an aspect of the language's grammar. To spell out this language-particular property in dynamical terms, a grammar potential function must be specified that contributes an attractor at a value of voicing x^0 , where $x^0 = [-\text{Voiced}]$. Assume that voicing is indexed with the parameter x of "glottal aperture," a tract-variable in Browman and Goldstein's dynamical representations. Then, as shown in (10), the grammar attractor appears at the right side of the x -axis, at some positive value of glottal aperture characteristic of voiceless obstruents. Below I explain how to derive the specific grammar potential function from basic assumptions.

(10) Potential function in a dynamical model of final devoicing



Intuitively, this is like the Optimality Theoretic formulation of a devoicing grammar “NOVOICEDCODA >> FAITH(Voice)” or the rule-based analog “[–Sonorant] [–Voiced] /___]°”. However, the dynamical formulation allows us to model the grammar’s interaction with context and ultimately derive phenomena like incomplete neutralization. Note that I do not examine the independent issue of why grammars develop properties like final devoicing. See Steriade (1997) for a proposal.

I now describe the grammar dynamics formally. As in any dynamical system, grammar dynamics is defined by a differential equation of the general form $dx/dt = G(x)$. Intuitively, this equation embodies the ‘dynamic law’ obeyed by the system. 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 (see section 7).

As a working hypothesis, I assume that the ‘tilted’ anharmonic oscillator provides a first approximation for the grammar dynamics: $G(x) = dx/dt = -k - x - x^3$. Several biological systems have been characterized by polynomial dynamics of degree three or less. For this and other reasons, our working hypothesis is not a bad assumption.

Given $-dV(x)/dx = dx/dt$ and $G(x) = -k - x - x^3$, we can compute by integration the potential for the grammar dynamics $V(x) = k*x + (-1/2) * (x^2) + (1/4) * (x^4)$ for some $k > 0$ and up to some constant term C which can be ignored as it does not affect the discussion or the qualitative results of the simulations. This $V(x)$ is the potential shown in the graph above.

The ultimate goal is to situate grammar in communicative context. We know that $G(x)$ has a stable point at the grammatically required value of $x = x^0$. We also know that the observed value of voicing is modulated by extra-grammatical parameters. Voicing is modulated by orthography, as shown in Ernestus and Baayen’s work (this volume), and by intentions as shown in Port and Crawford’s (1989) work. In what follows, I use intention as the extra-grammatical parameter, without loss of generality.

The basic fact of interest is that intentions can shift the preferred grammar modes. How can we formulate this in a principled way? The core idea to be fleshed out is that intentions contribute to the grammar an attractor corresponding to the intended form. The intention to communicate a lexeme with a final voiced consonant, in particular, is defined as a part of a

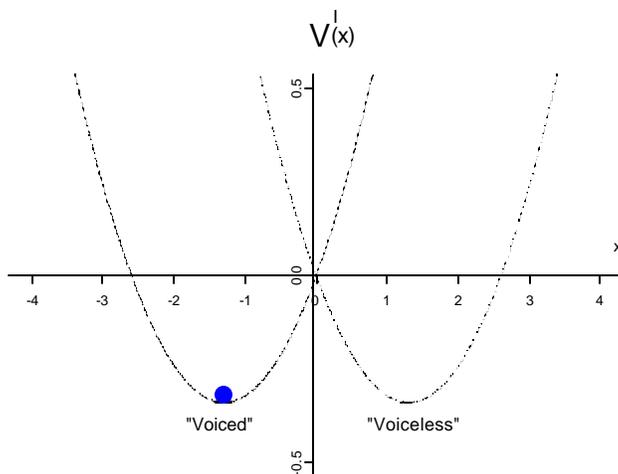
dynamics that attracts the order parameter toward the intended voicing. In turn, intentions are constrained by the grammar dynamics, namely, by how forms ‘should be produced’ in specific contexts. Overall, then, grammatical requirements sometimes compete and sometimes cooperate with variable environmental conditions (intentions). The phonetic output is the result of this combination of grammar dynamics and intentional dynamics. Incomplete neutralization will follow as a special case of this interaction.

6. Intentional dynamics

To situate grammar in communicative context, we need an appropriate dynamic formulation of intentions.

Informally, intentions are communicative goals. Let us assume a communicative act wherein the speaker’s goal is to convey the lexeme *Rad* ‘wheel’ as opposed to *Rat* ‘advice’. Intentional dynamics adds an attractor at the required value of voicing $\{-x^0, x^0\}$, where x^0 ‘=’ [-Voiced], $-x^0$ ‘=’ [+Voiced]. The potential $V(x)$ for these two values is shown below. Note that intentions are mutually exclusive. One can’t intend *Rad* and *Rat* at the same time – viz. the ball can only be in one of the two attractors.

(11) Dynamical model of “Voiced” and “Voiceless” intentions



I now describe the formal model for intentions. The dynamics of intentions in the context of a grammar G is modeled by the dynamical system $dx/dt = G(x) + I(x)$, following Schöner and Kelso (1988) on coordinated movement by humans. Intuitively, the ‘dynamic law’ obeyed by the combined system is given by a linear combination of the grammar dynamics $G(x)$ and the intentional dynamics $I(x)$.

$I(x)$ is the simplest function that adds an attractor at the (intentionally) required value of voicing. That is, $I(x) = \text{intent} * (x^{\text{REQ}} - x)$. In this function, ‘intent’ is a linear term representing the relative strength of the intentional contribution. The higher the value of ‘intent’, the stronger is the intention. The term x^{REQ} takes values from $\{-x^0, x^0\}$, that is, the values for glottal aperture corresponding to [+Voiced] and [-Voiced].

Given these assumptions, the contribution to the grammar dynamics that adds an attractor at the required value of voicing is given by the potentials shown above. To derive these

potentials, we start with $-dV(x)/dx = dx/dt = G(x) + \text{intent} * (x^{\text{REQ}} - x)$, and by basic calculus, we compute the part of the potential that corresponds to the intentional dynamics $V^I(x) = (1/2) * \text{intent} * (x^2) - \text{intent} * x^{\text{REQ}} * x$, 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. It is this $V^I(x)$ that is shown in the graph above.

I now sum up the essential ingredients of the proposal, in (12). There is a parameterization in terms of a ‘qualitative’ order parameter, the degree of voicing, and a ‘quantitative’ control parameter, the degree of intentional strength, shown in (12a,b) respectively. Order parameters describe the qualitative or macroscopic form of phonology and grammar principles refer to such parameters (see Gafos 2002 on gestural coordination relations). As seen in section 4, control parameters subject the order parameters to gradient drifts (the property of scaling). Beyond certain windows of variation, control parameters may lead to nonlinear jumps or bifurcations of the order parameters to different qualitative modes (the property of nonlinearity). In our example, the control parameter is intentional strength. As shown in (12c), there is also an ‘interface’, the hypothesized model relating these two parameters, $dx/dt = G(x) + \text{intent} * (x^{\text{REQ}} - x)$, where $G(x) = -k - x - x^3$. Crucially, however, this ‘interface’ does not translate symbols to continuous signals. Rather, it states a dynamic linkage, in the form of a testable relation, between a grammatical (order) parameter and a non-grammatical (control) parameter. The linkage is dynamic because the two parameters it relates are interdependent and changing quantities, as seen in section 3.

(12) Nonlinear dynamics as the linkage between the qualitative and the quantitative

- | | |
|--|---|
| a. x (degree of voicing) | order parameter; categorical/phonological |
| b. intent (degree of intentional strength) | control parameter; quantitative/scalar |
| c. $dx/dt = -k - x - x^3 + \text{intent} * (x^{\text{REQ}} - x)$ | the ‘interface’; the dynamic linkage between the order and control parameters |

In short, this is the core proposal of this paper, an alternative conception of the ‘phonetics-phonology interface’ where nonlinear dynamics offers a non-derivational way of relating qualitative and quantitative aspects of phonetics-phonology.

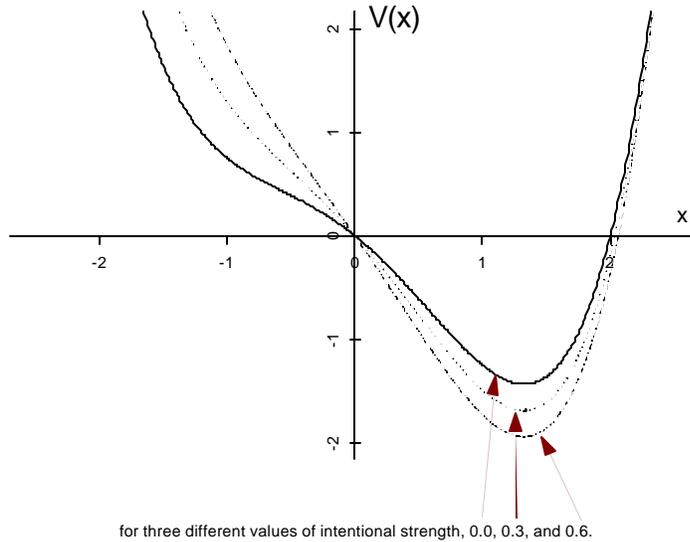
7. Simulations of grammar in varying intentional contexts

I now simulate the combined dynamics, grammar with intentional information. The parameters manipulated in the simulations are intention and its associated strength. Intention is categorically either Voiceless or Voiced, corresponding to the underlying value of the final obstruent in examples like *Rat*, *Rad*. Intentional strength is a scalar variable, which varies continuously in the interval $[0, 1]$. A value closer 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.

Consider first the case where the intention is a Voiceless obstruent, *Rat* ‘advice’. The intentionally required voicing value coincides with the grammatically prescribed value. They are

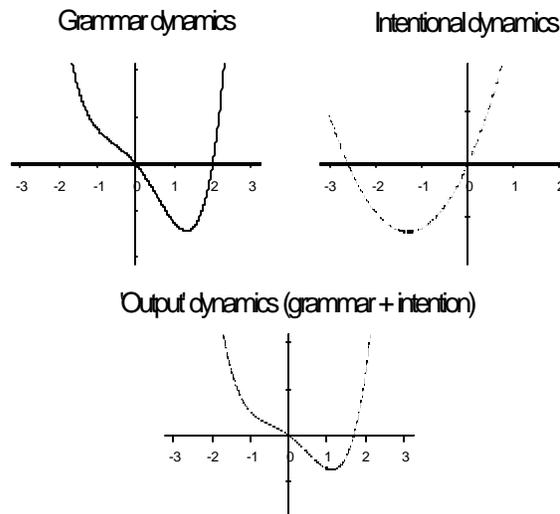
both Voiceless. In this case, then, we have cooperation of intentional requirements and grammar dynamics. As the figure below illustrates, there is no qualitative change in the resulting dynamics, indicated by the fact that the stable point remains fixed at the same value of x (x^0 '=[-Voiced]).

(13) Grammar dynamics as modified by intentional information [-Voiced]



Consider now the more interesting case where the intention is a Voiced obstruent, *Rad* 'wheel'. Here, the grammar dynamics contributes an attractor at the voiceless end of the x axis (the right side) and the Voiced intention contributes an attractor at the voiced end of the x axis (the left side). In this case, then, the intentionally required value does not coincide with the grammatically prescribed value. We have competition between grammar and intention. An instance of this competition is shown below.

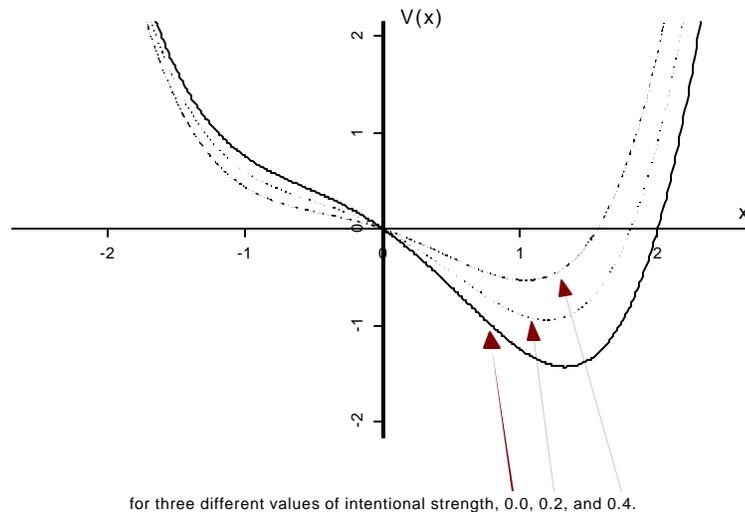
(12) Competition between grammar and intention, when intention is Voiced



The result of the competition schematized above is that the Voiceless attractor slowly drifts toward less extreme values. This scaling of the system's dynamics is shown more clearly in the figure below. It is observed that, as intentional strength increases, the potential is gradually pulled away from the [-Voiced] minimum toward more voicing. This is incomplete neutralization.

(14)

Grammar dynamics as modified by intentional information [+Voiced]



for three different values of intentional strength, 0.0, 0.2, and 0.4.

The effect of communicative context is directly captured in this model by the factor of intentional strength, and its effects on the dynamics. Overall, then, the two facts about neutralization, its incompleteness and its dependence on the communicative context, can be derived using basic concepts and tools of dynamics.

In simulations with this model not shown here, when the intentional strength for Voiced obstruents is increased beyond some relatively high value (> 0.78), the system changes discontinuously so that the only stable mode appears all the way at the other end. That is, the attractor is now at the Voiced end of the voicing continuum. The model then predicts a bifurcation, a qualitative change in the system's dynamics, as a result of a continuous increase in intentional strength. Indeed, if necessary, German speakers can produce Voiced obstruents as voiced in the neutralizing context (*Rad* as [ʁad]).

To sum up, the present model combines two seemingly incompatible insights from Ernestus and Baayen's (this volume) paper. The first is that "incomplete neutralization seems to be part and parcel of the grammar" (14). In the model, this is reflected in the way intentions parameterize the grammar. The second idea is that "incomplete neutralization may well be primarily a lexical effect" (13). This is reflected by identifying intentions with basic lexical forms. The intention for *Rat* is identified with an attractor at the voiceless end, whereas that for *Rad* with an attractor at the voiced end (of the order parameter, Voicing). As a consequence, intentions attract the order parameter toward the intended 'lexical' voicing. For voiced obstruents, specifically, incomplete neutralization follows.

8. Conclusion

The view of a phonological component preceding a phonetic implementation component is one way of expressing the intuition that phonetics-phonology is a system with qualitative and quantitative aspects (Ladd, this volume). However, it may not be the only way. A look at other complex systems may provide clues for alternative design methodologies. Given the preeminent view of language as a 'biological object' (Chomsky 2000), biological systems are the natural candidates. In theoretical biology (Waddington 1970, Pattee 1973), organisms described at the macroscopic level exhibit low-dimensional qualitative properties of considerable simplicity. At the microscopic level, the physicochemical processes of molecular biology are vastly detailed and continuous. Here, the temporal metaphor clearly fails. It does not make sense to say that the qualitative aspects of a living organism are related by precedence to their quantitative manifestations. The qualitative and quantitative coexist as two mutually dependent parts of a coherent whole.

Down to the more concrete level of analytical tools, the view of language as fundamentally biological suggests the use of the mathematics employed by leading physicists (Haken 1977) and biologists (Yates 1984) to study complex systems. As a small step in that direction, I hope to have shown some of the promise of nonlinear dynamics in providing a powerful formal method for addressing the issue behind the 'phonetics-phonology interface'. That is the issue of the relation between qualitative and quantitative aspects of phonetics-phonology. The proposal is that it is both necessary and promising to do away with the temporal metaphor of precedence between the qualitative and the quantitative, without losing sight of the essential distinction between the two. This leads to the alternative non-derivational conception of the term 'interface' as a dynamic linkage between the two interdependent aspects of a unified system.

Acknowledgment: Preparation of this paper was partly supported by a New York University Goddard Fellowship and Research Challenge Fund Grant, and an NIH Grant to Haskins Laboratories. The simulations reported were done in S-Plus 6 and Matlab 6.1. Thanks to Robert Port, Luigi Burzio, Paul Smolensky, Carol Fowler, Louis Goldstein, RenJ Kager and Wim Zonneveld for useful feedback, Janet Pierrehumbert, Donca Steriade, and Alice Turk for cogent comments and questions after presentation of this material at the Laboratory Phonology 8 meeting, and Arto Anttila, Richard Kayne and Anna Szabolsci at NYU for useful discussions. Any errors are my responsibility.

References

- Beer, Randall D. 1995. Computational and dynamical languages for autonomous agents. In Robert F. Port and Timothy van Gelder (eds.), *Mind as Motion*, MIT Press, Cambridge, Massachusetts, pp. 121-148.
- Boersma, Paul and Bruce Hayes. 2000. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32, 45-86.
- Browman, Catherine P. and Louis Goldstein. 1986. Towards an articulatory phonology. *Phonology* 3, 219-252.
- Browman, Catherine P. and Louis Goldstein. 1995. Dynamics and articulatory phonology. In Robert F. Port and Timothy van Gelder (eds.), *Mind as Motion*, MIT Press, Cambridge, Massachusetts, pp. 175-193.
- Cariani, Peter A. 1989. On the design of devices with emergent semantic functions. Ph.D.

- dissertation, State University of New York at Binghamton.
- Cedergren, Henrietta and David Sankoff. 1974. Variable rules: performance as a statistical reflection of competence. *Language* 50, 233-55.
- Charles-Luce, Jan. 1985. Word-final devoicing in German: effects of phonetic and sentential context. *Journal of Phonetics* 13, 309-324.
- Charles-Luce, Jan. 1997. Cognitive factors involved in preserving a phonemic contrast. *Language and Speech* 40, 229-248.
- Chomsky, Noam. 2000. *New horizons in the study of language and mind*. Cambridge University Press: Cambridge.
- Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Davidson, Lisa. 2001.
- Dinnsen, Daniel A. 1985. A re-examination of phonological neutralization. *Journal of Linguistics* 21, 265-279.
- Dinnsen, Daniel A. and Jan Charles-Luce. 1984. Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics* 12, 49-60.
- Eco, Umberto. 1984. *Semiotics and the philosophy of language*. Indiana University Press: Bloomington Indiana.
- Ernestus, Mirjam and Harald Baayen (this volume) The functionality of incomplete neutralization in Dutch: the case of past-tense formation.
- Fodor, J. A. and Z. W. Pylyshyn. 1981. How direct is visual perception? Some reflections on Gibson's 'ecological approach'. *Cognition* 9, 139-196.
- Fougeron, Cecile and Donca Steriade. Does deletion of French schwa lead to neutralization of lexical distinctions? *Proceedings of the 5th European conference on speech communication and technology* (University of Patras, Greece), vi, 943-946.
- Gafos, Adamantios. 2002. A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20, 269-337.
- Haken, Hermann. 1977. *Synergetics, An introduction*. Heidelberg: Springer-Verlag.
- Haken, Hermann. 1990. Synergetics as a tool for the conceptualization and mathematization of cognition and behavior – How far can we go?. In H. Haken and M. Stadler (eds.), *Synergetics of Cognition* (Springer Series in Synergetics, Vol. 45), pp. 2-31, Heidelberg, Germany: Springer-Verlag.
- Harnad, Stevan. 1990. The symbol grounding problem. *Physica D* 42, 335-346.
- Hintzman, D. 1986. "Scheme abstraction" in a multiple-trace memory model. *Psychological Review* 93, 411-428.
- Kelso, Scott J.A. 1984. Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology* 15, R1000-R1004.
- Kelso, Scott J.A. 1995. *Dynamic patterns: the self-organization of brain and behavior*. Cambridge, MA: MIT Press.
- Kelso, Scott J.A., Mingzhou Ding, and Gregor Sch`ner. 1992. Dynamic pattern formation: a primer. In A. Baskin and J. Mittenhal (eds.), *Principles of organization in organisms*, pp. 397-439. Sante Fe, NM: Santa Fe Institute; and Reading, MA: Addison-Wesley.
- Kingston, John and Mary E. Beckman. 1990. *Papers in Laboratory Phonology I: between the grammar and the physics of speech*. Cambridge University Press: Cambridge.
- Kingston, John and Randy Diehl. 1994. Phonetic knowledge. *Language* 70, 419-54.
- Ladd, Robert D. (this volume) "Contrastive phones" in surface representation.

- Lass, Roger. 1997. *Historical linguistics and sound change*. Cambridge: Cambridge University Press.
- Lieberman, Mark Y. 1983. In favor of some uncommon approaches to the study of speech. In Peter F. MacNeilage (ed.), *The Production of Speech*, New York: Springer-Verlag.
- Miller, Joanne. 1981. Effect of speaking rate on segmental distinctions. In *Perspectives on the study of speech*, pp. 39-74.
- Pattee, Howard H. 1973. The physical basis and origin of hierarchical control. In H. H. Pattee (ed.), *Hierarchy Theory: The Challenge from Complex Systems*, pp. 71-108, New York: Braziller.
- Petitot-Cocorda, Jean. 1985. *Les catastrophes de la parole*. De Roman Jakobson B RenJ Thom. Paris: Maloine.
- Pierrehumbert, Janet. 2001. Exemplar dynamics: word-frequency, lenition and contrast. . In Bybee, J. and P. Hopper (eds.), *Frequency effects and the emergence of linguistic structure*, pp. 137-157, John Benjamins, Amsterdam.
- Pierrehumbert, Janet. 2002. Word-specific phonetics. In C. Gussenhoven and N. Warner (eds), *Laboratory Phonology VII*, Mouton de Gruyter, Berlin.
- Port, Robert and Michael L. O'Dell. 1985. Neutralization of syllable-final voicing in German. *Journal of Phonetics* 13, 455-471.
- Port, Robert and P. Crawford. 1989. Incomplete neutralization and pragmatics in German. *Journal of Phonetics* 17, 2257-282.
- Port, Robert and Timothy van Gelder (eds.). 1995. *Mind as motion: Explorations in the dynamics of cognition*. Bradford books, MIT Press.
- Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*, unpublished manuscript, Rutgers University and University of Colorado, Boulder.
- Saltzman, Elliot. 1995. Dynamics and coordinate systems in skilled sensorimotor activity. In Port and van Gelder (1995), pp. 149-174.
- Sankoff, David. Variable rules. In *Sociolinguistics: an international handbook of the science of language and society*, pp. 984-97. Berlin, New York: Walter de Gruyter.
- Schöner, Gregor and J. A. Scott Kelso. 1988. A dynamic pattern theory of behavioral change. *Journal of Theoretical Biology* 135, 501-524.
- Slowiaczek Louisa M. and Daniel A. Dinnsen. 1985. On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13, 325-341.
- Smolensky, Paul. 1988. On the proper treatment of connectionism. *Behavioral and Brain Sciences* 11, 1-74.
- Sommerstein, Alan H. 1977. *Modern phonology*. Edward Arnold: London.
- Steriade, Donca. 1997. Phonetics in phonology: the case of Laryngeal neutralization. Los Angeles: UCLA ms. [Published in UCLA Working Papers in Phonology.]
- Stevens, Kenneth N. 1972. The quantal nature of speech: evidence from articulatory-acoustic data. In *Human Communication: A Unified View*, ed. E. David and P. Denes, 51-66. New York: McGraw-Hill.
- Stevens, Kenneth N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17, 3-45.
- Summerfield, Quentin. 1981. Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance* 7, 1074-95.
- Titze, I. R. 1988. The physics of small amplitude oscillation of the vocal folds. *Journal of the Acoustical Society of America* 83, 1526-1552.

- Thompson, J. M. T., and H. B. Stewart. 2002. *Nonlinear dynamics and chaos*. John Wiley & Sons. [2nd edition.]
- Trubetzkoy, N.S. 1969. *Principles of phonology*. Berkeley and Los Angeles: University of California Press. [translation by C. Baltaxe of *Grundzüge der Phonologie* (Goettingen: Vandenhoeck and Ruprecht, 1958)]
- Tuller, Betty, Pamela Case, Mingzhou Ding, and J. A. Scott Kelso. 1994. The nonlinear dynamics of speech categorization. *Journal of Experimental Psychology* 20 (1), 3-16.
- Turvey, M. T. 1990. Coordination. *American Psychologist* 45, 938-953.
- Waddington, C. H. 1970 (ed.). *Towards a theoretical biology* (Vol. 3). Aldine Publishing Company: Chicago.
- Winfree, Arthur T. 1980. *The geometry of biological time*. Springer-Verlag, New York.
- Wing, Alan M., and Peter J. Beek. 2002. Movement timing: a tutorial. In *Attention and Performance XIX: Common mechanisms in perception and action*, eds. Wolfgang Prinz and Bernhard Hommel, pp. 202-226, Oxford University Press.
- Yates, F. E (ed.). 1984. *Self-organizing systems: The emergence of order*. New York: Plenum Press.
-
-