Back to the past tense in English*

Louis Goldstein
University of Southern California
Haskins Laboratories

I. Introduction

The regular alternation of the English Past tense is commonly taught in introductory linguistics and phonology classes (the latter of which was one of the origins of my friendship with Judy Aissen). Yet at the same time, this relatively simple alternation has provided important challenges and has been a touchstone for novel theoretical approaches and techniques in phonology (from extrinsic rule ordering to connectionism to optimality theory to wug-testing). Here we re-consider (yet again) the allomorphy of the English regular past-tense, this time within a dynamical, articulatory framework in which (i) the combinatorial units of phonology are abstract vocal constriction actions, (ii) held together by dynamical principles of coordination, and in which (iii) phonological grammars explicitly characterize the stable patterns of coordination employed by a language, and the possible context-governed shifts in pattern that the language exhibits.

The last thirty years has seen development of a dynamical approach to cognition, in which a single formal language (nonlinear dynamics) is used to model both the qualitative (discrete) and quantitative (continuous) aspects aspects of complex systems, including those underlying cognition and action (Kelso, Scholz, & Schöner, 1986; Kugler & Turvey, 1987; Turvey, 1990; Kelso, 1995). The key idea is that qualitative cognitive categories can be understood as the stable (low-dimensional) modes of a nonlinear dynamical system, defined as a differential equation in a continuous state space. Due to the nonlinearities in such systems, changing one or more of the differential equation's parameter values ("control" parameters) in a continuous fashion can lead to lawful shifts in the qualitative (categorical) state of the system. A classic example of this (Haken, Kelso & Bunz, 1985) can be observed in how oscillating limbs are coordinated with each other. When the rate of oscillation is increased, the limbs will exhibit a qualitative shift from the two limbs moving 180 degrees out of phase with one another to moving in phase.

The development of Articulatory Phonology (1990a,b; 1992a; 1995) was an early attempt to apply these ideas in the domain of phonological *representation*. By defining the primitives of phonology to be speech gestures, each of which is modeled by a dynamical system that regulates the motion of articulators in the formation of constrictions (see below), it became possible to *lawfully* relate qualitative, phonological representations (in terms of gestures and their coordination in time) to continuous physical movement. (Browman & Goldstein, 1995). An important consequence of this lawful relation is that it is possible to use articulatory (and acoustic) data to directly test hypotheses about phonological structure (Gafos & Goldstein, to appear). So for example, Browman & Goldstein (1990c) and Zsiga (1995) showed that the output of certain (post-lexical) assimilation processes was a phonological representation that included all of the gestures of the non-assimilated forms, but in a different temporal relation (more overlap in time). No gestures were deleted or changed.

More recently, dynamical systems have been applied more widely in phonology, to include processes of phonological planning (Nam, 2007; Gafos & Kirov, 2009; Nam, Goldstein & Saltzman, 2009), to interactions among phonological units during planning of the sort that result in speech errors (Goldstein, Pouplier, Chen, Saltzman & Byrd, 2007), and to certain types of sound change (Parrell, submitted; Hsieh, 2010). These examples show how the nonlinear properties of the relevant dynamical systems can cause qualitative shifts as a function of changes in a (continuous) control parameter. For example, Parrell (submitted) has investigated an ongoing sound change in Western Andalusian Spanish, in which the h-stop sequences that can appear across syllable boundaries (due to lenition of coda /s/ to [h]) are (qualitatively) reorganized into aspirated stops. As Torreira (2007a,b) has argued this can be viewed as a shift in coordination of a glottal abduction gesture and an oral constriction gesture from a sequential (anti-phase) mode of coordination mode to a (dynamically more stable) synchronous (in-phase) mode. Parrell's experiment provides quantitative evidence for the nonlinear dynamical system analysis, including dependence on speaking rate (the shift occurs more frequently as speaking rate increases), and increased variability of intergestural coordination for the tokens that lie between the two stable modes.

The most general application of these ideas to phonology can be found in the work of Gafos (e,g., Gafos, 2006; Gafos & Benus, 2006), in which the synchronic phonological grammar itself is a nonlinear dynamical system. Individual constraints can be formalized as nonlinear dynamical equations whose modes correspond to the preferred value or values of gesture targets or of intergestural coordination parameters. Through the dynamical interaction of multiple constraints, and through modulation of a constraint's control parameters by the current state of the system, contextually appropriate shifts in targets are selected. Gafos & Benus (2006b) account for experimental data on transparency in Hungarian vowel harmony using this kind of grammar-dynamical system.

Analysis of allomorphy in the regular English past tense is particularly interesting to discuss in light of these theoretical developments, because it shows how understanding of these alternations can result from maintaining *both* a dynamical theory of representation (gestures and their coordination) and also a dynamical theory of grammar. As we will see, one of the alternates is completely predictable from the hypothesis that phonological units are gestures and that affixation operates by specifying a dynamical coordination relation between stem and affix. The other very nicely demonstrates the need for some nonlinear system that effectively that governs the *selection* of coordination relations, though the formal development of that system will be left for the future.

II. Dynamical representation in phonology: Articulatory phonology

Articulatory phonology (Browman & Goldstein, 1992a; 1995) hypothesizes that the combinatorial units of phonology are abstract, dynamical representations of speech production actions, called gestures. These gestural actions have as their goals the formation & release of constrictions within the vocal tract. The abstract language in which gestural representations are defined, dynamical systems, is different from the abstract symbolic system employed in classical phonological representations, but it is nonetheless abstract in two relevant senses, roughly, temporal and spatial. The temporal sense follows from the fundamental insight of dynamical systems, as conceived by Newton. The insight is that while we typically observe continuous change when we look at the world, it is possible to find events during which the continuous change is the lawful consequence of mathematical relationships (the dynamical system) whose parameters do *not* change over the time span of the event. For example, a falling body is constantly accelerating from when it is dropped to when it hits the ground. But the equation of motion that describes that motion is unchanging during that event. Likewise, when we observe the articulators of the vocal tract during speech, they are constantly in motion and their velocity is changing. But research over the last 25 years has shown that (to a first approximation) the motion and velocity change can be modeled as the lawful consequence of the individual gestures' dynamical systems (their equation of motion). Importantly each gesture is an event that is active in the vocal tract for some fixed interval of time, during which its equation of motion (including the specification of its parameter values) is fixed and unchanging. The temporal discreteness of gestures is one of the properties that allows them to function as combinatorial units of phonology.

Gesture dynamics are also abstract spatially. Each gesture is defined as a task (Saltzman & Munhall, 1989) that forms (or releases) a constriction within the vocal tract. The equation of motion for these tasks do not govern the positions of individual vocal tract articulators, but rather the abstract quantities of the constriction size (and location), formed with a given "constricting device" in the vocal tract (Lips, Tongue Tip (TT), Tongue Body (TB), Glottis (GLO), and Velum(VEL)). The same constriction can (and is) achieved with different combinations of articulator motions in different contexts. Such differences are lawful consequences of the individual gestures' dynamical specifications and their patterning in time, as formalized within the task dynamics model (Saltzman & Munhall, 1989) that underlies Articulatory Phonology representations. Thus, each gesture has a contextually invariant parameter specification, which allows them to function as combinatorial phonological units. Contrasting gestures, can differ in parameter specifications. For example, /t/ and /s/ are represented by tongue tip gestures that differ in their values for the constriction degree (CD) of the tongue tip (TT).

Putting together the temporal and spatial abstractness of gestures, it is possible to display the dynamical control for a given phonological form over time in a *gestural score*. Each box in the gestural score represents the interval of time (along the horizontal axis) during which that gesture's dynamical equation of motion is *active*. For example, the gestural score in Fig. 1 for the utterance "nap" shows the activation intervals for the gestures of each of the constricting devices composing this form. For example, at left (beginning of the utterance), there are gestures corresponding to the initial /n/: a wide velic constriction (*VEL-wide*) that opens the aperture to the nasal cavity ,and a closure gesture of the tongue tip (*TT-clo*) followed immediately by a release of the tongue tip constriction (*TT-rel*). Closures, as well as consonant gestures of other constriction degrees (e.g., those appropriate for fricatives (*crit*), and liquids/glides (*narrow*)), are typically followed by actively-controlled releases. The gesture of the tongue body (*TB-wide*) forms the constriction for the vowel. Note that the gestures overlap in time. This is perhaps not surprising for the gestures that compose individual segments (e.g, *VEL-wide* and *TT-clo* for /n/), but it is also found for gestures that compose successive segments (*TT-clo*, *TB-wide*). *LA-clo* and *LA-rel* are lip closure and release gestures (LA=Lip Aperture).

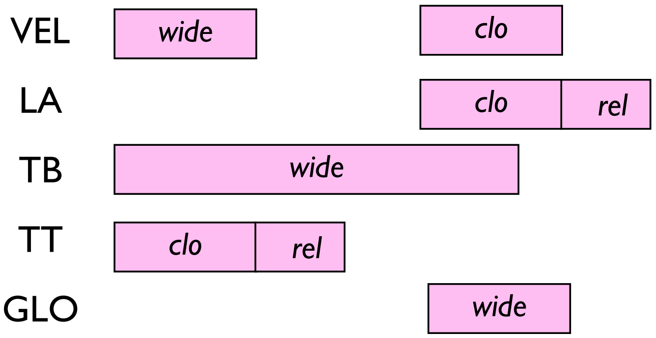

Figure 1. Gestural Score for “nap”

A gestural score displays activation of gestures over continuous time. However, the actual time of activation of individual gestures varies considerably as a function of the prosodic context and speaking rate, and so the gestural score in Fig. 1 depicts a canonical production of “nap.” Due to this variation, the gestural score itself is not an abstract, invariant phonological representation of temporal organization. Rather, it is hypothesized that there is some dynamical representation of the coordination of gestures in time that is *characteristic* of that particular utterance from which the activation times in particular contexts can emerge lawfully. Recent work (e.g., Saltzman, Nam, Krivokapic & Goldstein, 2008) has developed a dynamical model of speech gesture planning, in which the coordination of gestures is modeled by assigning each gesture to an internal oscillator or clock, that is responsible for triggering it in real time, and by specifying dynamical coupling relations between pairs of gesture clocks. The ensemble of multiply-coupled clocks is a nonlinear dynamical system.

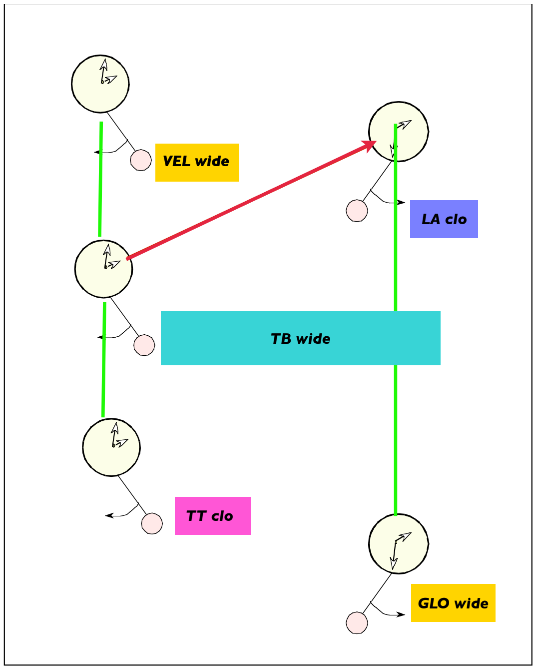

Figure 2. Coupling graph for “nap”

Research over the last 25 years on coordination of rhythmic movements of multiple limbs (Haken, Kelso & Bunz, 1985; Turvey, 1990) has revealed that there are two qualitatively distinct modes of coupling oscillators that are intrinsically stable and are readily performed without requiring any learning—*in-phase* (the most stable and most accessible) and *anti-phase* (180° degrees out of phase). The planning model hypothesizes that these discrete modes are employed in coupling gestures' clocks, and that they form the basis for a dynamical model of syllable structure (Goldstein, Saltzman & Byrd, 2006; Nam, 2007; Nam et al., 2009). Onset consonant gesture clocks are hypothesized to be coupled in-phase with the vowel gesture clock. Since the clocks trigger gesture activation, onset consonant gestures and vowel gestures are triggered synchronously. Evidence for this synchrony in triggering can be found kinematic data showing rough synchrony in the onset of articulator movements of the onset consonant and the vowel (Löfqvist & Gracco, 1999). The accessibility of this mode can also be used to account for the typological preference for onsets over codas (Goldstein et al, 2006). The clock of a coda consonant is hypothesized to be coupled anti-phase with the vowel gesture clock, and the triggering of the coda consonant gesture will be delayed by 180° from the triggering of the vowel. (The amount of time corresponding to 180° will depend on the frequency of the clocks, which will depend on speaking rate). Thus, in-phase coupling is used to model observed patterns of *synchronous* gestural coordination, while anti-phase is used to model *sequential* coordination. Fig. 2 shows the coupling relations among the clocks for the utterance "nap." (The clocks for the release gestures have been left out to keep the figure simple, as is the *VEL-clo* gesture for /p/).

The display in Fig. 2 is a graph, in which the nodes correspond to gestures and the edges correspond to coupling relations. Such *coupling graphs* can be taken to be part of the phonological representation of utterances, and contrast among graphs can be defined either in the set of nodes, or in the topology of the links. In-phase coupling is represented by green edges without arrows, while anti-phase coupling is by red edges and arrow heads pointing to the gesture that is triggered later. Specification of order of triggering is required for any coupling relation that is not in-phase (ie., synchronous). Within the model, planning and production of a form begins with the graph. During planning, oscillators are set in motion at random initial phases, but are coupled according to the graph specifications. Over time, the coupling causes the clocks to settle into stable phase relations, and once they are stabilized, the clocks trigger their associated gestures.

While in-phase and anti-phase coupling can be performed without any learning, other *eccentric* coupling modes (arbitrary relative phases) can be learned, in order to perform more difficult coordination tasks, such as juggling or drumming. The syllable structure model hypothesizes that eccentric coupling is used to coordinate consonant gestures in an onset or a coda a cluster (and in some cases across syllables). This can be one of the reasons that clusters are acquired relatively late by children (Nam et al, 2009), and why they are relatively marked typologically. The particular eccentric coupling employed may differ from language to language (Zsiga, 2000; Yanagawa, 2006; Goldstein, Chitoran & Selkirk, 2007). Onset clusters have been investigated extensively and have been shown, across a wide variety of languages, to exhibit a competitive

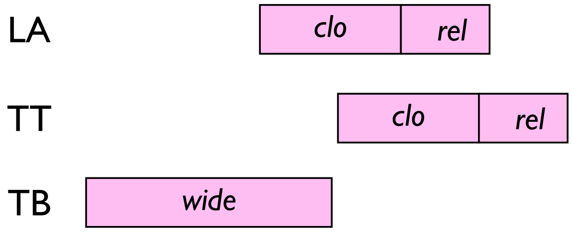

Figure 3. Gestural score for "apt"

coupling topology in which all onset Cs are in-phase coupled with the vowel, but also coupled to one another either anti-phase or eccentrically (e.g., Browman & Goldstein, 2000; Goldstein, Chitoran & Selkirk, 2007; Nam, 2007; Saltzman et al, 2008; Hermes, Grice, Mücke, & Niemann, 2008; Shaw, Gafos, Hoole & Zeroual, 2009; Marin & Pouplier, 2010). However, our focus in this paper is on the less-studied coda clusters. In English, at least, there is evidence (Byrd, 1995; Honorof & Browman, 1995; Marin & Pouplier, 2010) that the first coda consonant (C_1) is coupled anti-phase to the vowel and that a following C2 is eccentrically coupled to the first. The relative timing of consonant gestures in coda clusters in English can be illustrated by the gestural score in Fig. 3 for the word "apt" (oral and velic gestures are left out for simplicity). C2 (TT) activation is triggered roughly halfway between the onset of C1 (LA) and the time the that C1 is deactivated and its release begins. That is, C2 *overlaps* C1 by half of its duration. This coordination relation between *C1-clo* and *C2-clo* we will refer to as "shingled" or "semi-overlapped." In the TaDA model (see next paragraph), consonant gestures are active for 60° of their clock's cycle. Therefore, shingled coordination can be achieved by an eccentric coupling of 30°, which is employed for CLO-CLO coupling in coda.

This gestural score for "apt" can be produced by the coupling graph in Fig. 4. The eccentric coupling relation between the *LIP-clo* and *TT-clo* is represented in blue. Note that the *clo* and *rel* gestures of a given constriction (LIP, TT) are connected with a *black* headed arrow. This eccentric coupling relation between *clo* and *rel* is a special one, with the *rel* being triggered automatically when the clo is deactivated ("abutting" activations). This coordination is not the same as the shingled coordination of two Cs in a cluster.

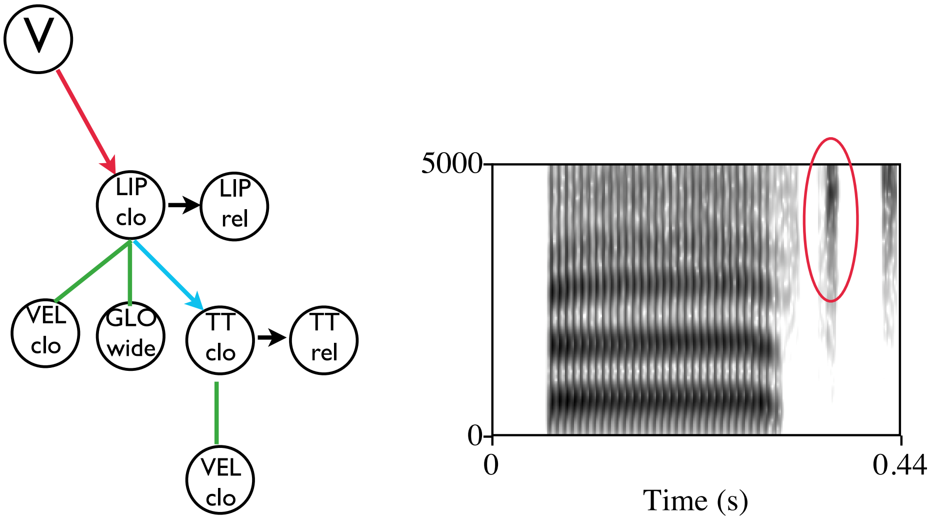

Figure 4. Coupling graph and spectrogram of TaDA output for "apt"

The coupling graph model of syllable structure, the coupled oscillator model of planning, the task-dynamic model of articulator coordination, and an articulator-based vocal tract shaping model are all integrated into the computational system called TaDA (Nam, Goldstein, Saltzman & Byrd, 2004). The output of TaDA (time-varying constrictions and vocal tract shapes) can be input into the HLsyn model (Hanson & Stevens, 2002), to generate simulated aerodynamics and output sound. The spectrogram in Fig. 4 shows the audio output when the coupling graph is input to TaDA. Note that the /p/ is acoustically released (the release burst is circled in the spectrogram). Shingled coordination will typically result in release of heterorganic CC clusters, although when speech rate in the model is increased, the release can disappear. This is consistent with the fact that C1 in such clusters in English can produced with audible releases or without them.

III. Gestural representation of English regular past-tense allomorphs

As is well known, the regular English past-tense suffix takes three phonological forms: /-d/ in most contexts; /-t/ following voiceless consonants (except /t/); and /-Vd/ following coronal (oral) stops, where V has been represented as /ə / or /ɨ/. By employing dynamical gestural representations, however, it is possible to assign the suffix a *single* representation, namely a (sub)-graph consisting of a Tongue Tip (TT) closure gesture, coupled with a velic closure (active closure of the soft palate to insure that the closure is non-nasal), and with a TT release gesture. The varying allomorphs result from principles that determine how that subgraph is coordinated with the coupling graph of the stem, and from the articulatory, aerodynamic, and acoustic consequences of the resulting composite graph. The main generalization for coupling of the morphological suffix is this:

(1) Couple the TT closure gesture of the suffix graph like a coda consonant:
  (a) anti-phase to the V, if the stem ends in V
  (b) eccentric (30°) to a stem-final C to produce semi-overlapped coordination (CLO-CLO coupling)

A. Voicing Alternation: /-d/ vs. /-t/

The /-d/ and /-t/ allomorphs can be analyzed as resulting from *identical* coordination patterns. The difference in voicing is hypothesized to result from the aerodynamic *consequences* of coordinating the hypothesized suffix subgraph with stem graphs that differentially affect the probability that voicing will be produced. Note that there is nothing in the representation of the suffix graph itself that directly controls voicing (no glottal abduction gesture that would inhibit voicing, no larynx lowering or oral cavity expansion gestures that would enable voicing to continue during a closure interval). In that sense the graph is bit like a traditional archiphoneme (Trubetzkoy, 1939). Stem graphs that end in oral closure gestures and with glottal abduction gestures will inhibit the likelihood that the closure and release caused by the suffix gestures will be produced with observable voicing. These differing consequences can be observed by inputting the relevant coupling graphs to the TaDA/HLsyn model, which computes the articulatory, aerodynamic and acoustic consequences of the pattern of constriction gestures in the vocal tract. In evaluating the output, it is important to note that there is no quantitative, acoustic data (to my knowledge) on the voicing properties of some of these forms as produced by native speakers. So for example, while a form like "nabbed" is analyzed as having the /-d/ suffix, it is unclear how frequently the final coronal closure and release is actually produced with periodic vibration, given that (phonologically and perceptually) voiced stops in English are often not produced with such vibration. So the point of the examples presented below is to demonstrate that attachment of an invariant suffix to a stem can have a *variety* of voicing consequences, consistent with the aerodynamic properties of the stem gestures, and consistent with the distribution of the traditional allomorphs. Detailed comparison of the model's output with that of native speakers will not be attempted here.

First we consider the case of a stem-final vowel. According to (1a), the TT gesture is coupled to the vowel in anti-phase mode. Fig. 5 shows the coupling graph for the past tense of the verb "sew". The shaded graph nodes correspond to subgraph for the suffix. This context is optimally favorable for voicing, as there is high airflow and phonation during the vowel, Spectrogram of the output from TaDA on the right shows that the TT closure and its release are both voiced, as is consistent with the traditional description of this allomorph as being voiced (/d/).

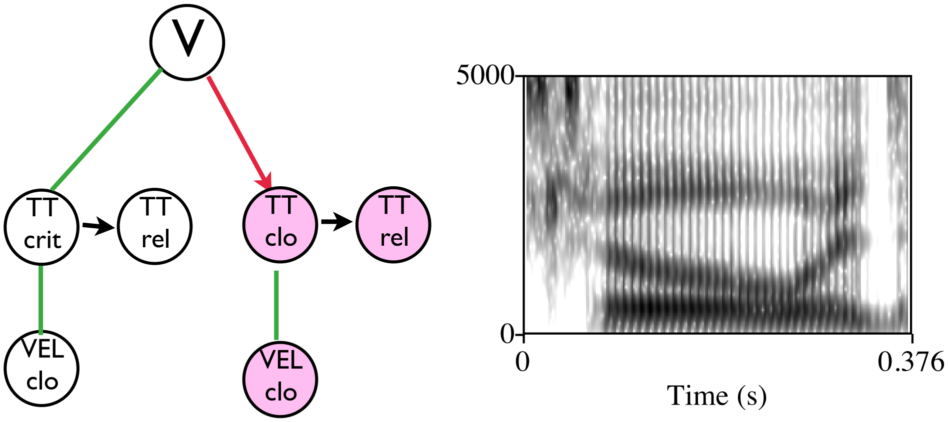

Figure 5. Coupling graph and spectrogram of TaDA output for "sewed"

Now, let us consider a stem-final voiced C, for example the word "dimmed." Here, according to (1b), the TT gesture of the suffix should be shingled to the final C of the stem. This CLO-CLO coupling is eccentric with 30 degree phase offset. Fig. 6 shows this coupling graph for "dimmed," along with the spectrogram of the acoustic output from TaDA for this coupling graph. The velic opening (*VEL-wide*) gesture coordinated with the lip closure (*LA-clo*) of the /m/ results in nasalization and also contributes to favorable conditions for voicing. Voicing can be observed through the nasal closure, and even into oral closure, and the release is voiced. Some devoicing occurs during closure, the aerodynamic result of a complex of events including: onset of velic closure, release of lip closure, and formation of TT closure.

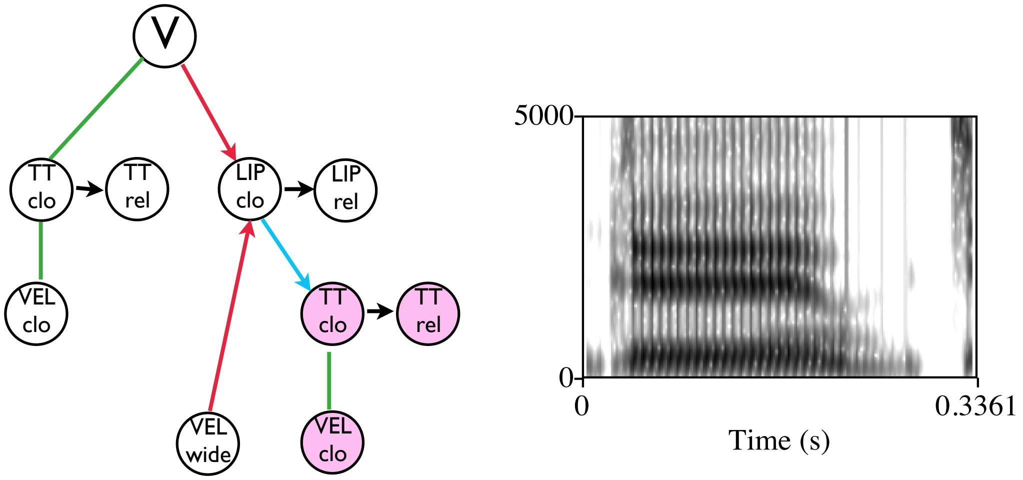

Figure 6. Coupling graph and spectrogram of TaDA output for “dimmed”

We can contrast the output voicing in this case with the results when the suffix graph is coordinated to the stem for “nap,” as in Fig.7. The *TT-clo* gesture of the suffix is coupled to the *LA-clo* gesture of the stem, exactly as in “dimmed.” Now however, because of the glottal abduction gesture (*GLO-wide*) coupled to the lip closure, and because there is velic closure instead of velic opening (as in “dimmed”), conditions are maximally *un*favorable for voicing, and we can see that *no* voicing is found during the lip closure, its release, the tongue tip closure, or its release. This state of affairs is consistent with the description of the allomorph of the past tense as /-t/ in this condition. Note, however, the the phonological form of the suffix (the coupling subgraph) and its coupling to the stem are both identical in this case and in “dimmed.” The difference in output acoustics is caused by the aerodynamic interactions of stem and suffix gestures.

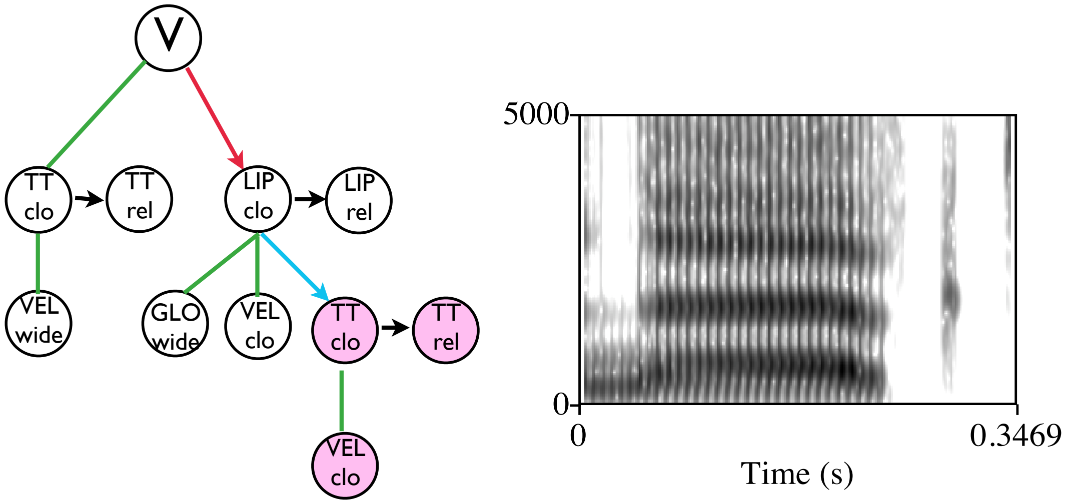

Figure 7. Coupling graph and spectrogram of TaDA output for “napped”

Finally, we examine the case of a stem-final voiced stop, as in “nabbed.” The expectation here is that the aerodynamic consequences of coupling the suffix graph to the stem graph would be a voiced coronal stop, if it is to be consistent with the /-d/ allopmorph. As can be seen in Fig.8 however, while the closure for the /b/ is largely voiced, its release, and the coronal stop closure and its release are all voiceless. Presumably this is due to the aerodynamic consequences of the extended closure interval caused by the multiple stops. Although there is a short release between the lip and tongue tip closures, it is very short and perhaps does not afford a return to pressure conditions that allow glottal vibration. However, the output sounds like a good exemplar of “nabbed.” While there has been no systematic acoustic study of voicing in these contexts, a spectrogram of one natural production of “nabbed” is shown in Fig. 9. It is strikingly like that of the model—there is voicing during the lip closure, but everything after that is voiceless.

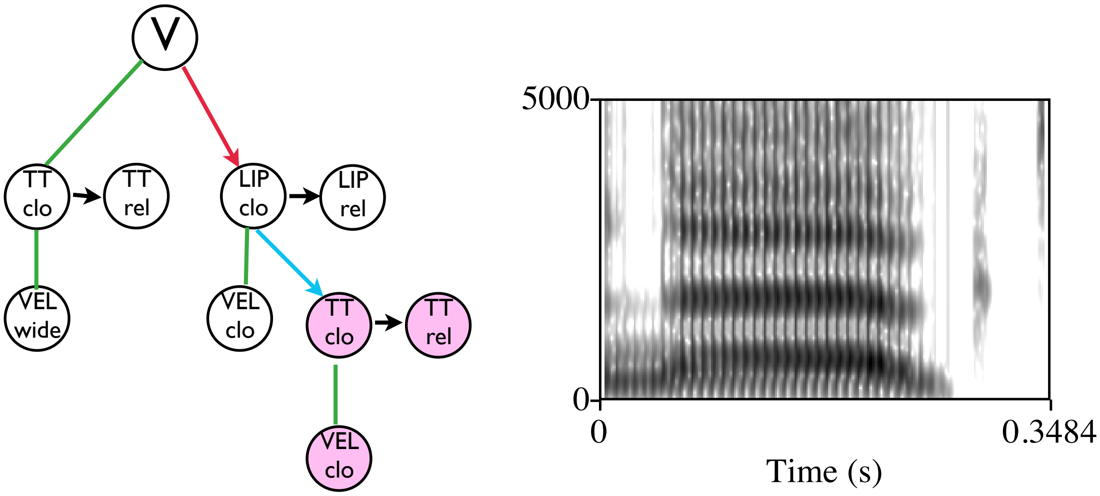

Figure 8. Coupling graph and spectrogram of TaDA output for “nabbed”

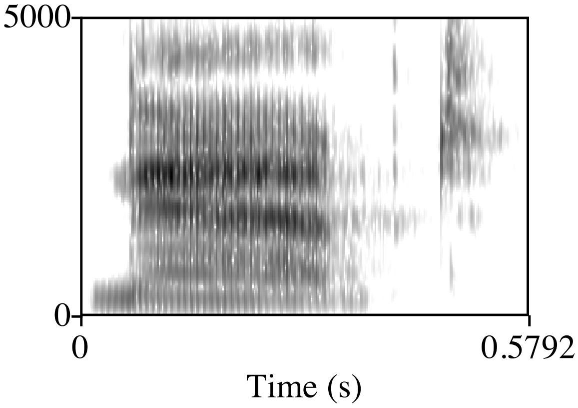

Figure 9. Spectrogram of natural token of “nabbed”

To summarize then, we hypothesized that the past tense could be represented as an invariant coupling sub-graph, composed of *TT-clo*, *VEL-clo* and *TT-rel* gestures. Coupling the *TT-clo* gesture to the stem in a manner consistent with its being treated as an (additional) coda consonant on the stem-final syllable leads to acoustic output that varies with the nature of the stem-final consonant, with a distribution consistent with the traditional statement of /-d/ vs /-t/ allomorphy. Importantly, however, if we take the coupling graphs to be the (syllable-level) phonological representation of the these forms, we note that there is no phonological alternation here, and no allomorphy. The abstract phonological representation of the suffix and its coupling is invariant. Output variability is due to the aerodynamic consequences of the stem gestures.

B. Syllabicity alternation: /-d/ vs. /-Vd/

When the past-tense subgraph is coordinated to stems ending in /t/ or /d/ using the coordination principle in (1), then the resulting form does not have a reduced vowel (/ə / or /ɨ/) between the stem-final coronal and the suffix coronal. (The actual model output is discussed in detail below). So allomorphy needs to be analyzed as resulting from a change in the stem suffix or in the operative coordination principle (or both). Because a grammar in Articulatory Phonology operates on coordination relations, it is possible for the grammar to account for the systematic occurrence of a vowel (particularly a reduced one) in some context, without requiring that the vowel results from deploying a vowel gesture with a constriction task (or target). Rather, the observed vowel interval can emerge from coordination of two consonant gestures (C_1C_2) in a temporal relation such that C_1 is completely released before C_2 begins to be formed. This will leave a temporal "gap" between the constrictions whose vocal tract shape is not determined by a vowel target, but rather by the articulator motions resulting from the release C_1 and by movement of individual articulators to their rest postures when they are not being controlled by any active gesture. The formalization of this type of "temporal-only" grammatical specification using coordination constraints plays a key part in Gafos' analysis of template phonology in Moroccan Arabic (Gafos, 2002).

Since a 'targetless' temporal gap will be relatively unconstricted and variable (its vocal tract shape will largely be determined as a function of flanking consonants and vowels), this kind of (temporal-only) control seems a plausible model for reduced vowels, such as the one that appears in the /-Vd/ plural allomorph. Browman & Goldstein (1992b) tested the hypothesis that (lexical) schwa vowels in English were 'targetless' in this sense, but they had to reject the hypothesis. There was evidence that there was a specific constriction target associated with those schwa vowels. However, the reduced vowel of the plural affix (e.g., "roses") have been shown (Flemming & Johnson, 2007) to be acoustically different from the lexical schwa vowels in final syllables (e.g., "Rosa's"), as is consistent with the transcription of the plural (and past tense) affixes as [ɨ], but the lexical reduced vowel as [ə], a tradition going back to Trager & Smith (1951). Flemming & Johnson showed that the plural affix vowel is higher (has a lower F1) than the one found in final lexical schwas. Since the tongue shape in these vowels is also relatively fronted (F2 in their study ranges from 1750 to 2200 Hz), the acoustics seem consistent with a largely coronal fricative configuration, but with a lowered tongue tip. Thus, these reduced vowels appear to be good candidates for purely temporal gaps between release of one coronal and formation of the next, and given the similar transcriptions for the past-tense suffix, and given the similar coronal context, these appear to be good candidates as well.

The targetless hypothesis for the past tense affix was tested by Smorodinsky (2001). She collected kinematic data from the tongue, jaw, and lips (using electroarticulography, EMA) while speakers read utterances like those in (2), with near-minimal pairs of lexical vs. past-tense affix vowels:

(2) “If Cheetah’d even known” (lexical)
“If cheated even once” (affix)

The full vowels surrounding the reduced vowel were symmetric (same vowel preceding and following) and pairs were constructed using all the full monophthongs of English. She found that the position of the tongue receivers during the affix vowels were more correlated with the positions of the flanking full vowels than was the case for the lexical vowels. This was taken as evidence that the affix vowels could be produced with no active tongue body control--the shape during those vowels emerging from the shape of the flanking vowels and consonants. While these results are consistent with the targetless analysis, there is a methodological problem using the EMA device for this purpose, as it can only measure the position of the front part of the tongue, and for schwa-type vowels, it is possible that there are significant tongue-shaping events occurring posterior to the tool’s reach. Recently, these results have been replicated using real-time MRI (Lammert, Goldstein & Narayanan, 2010), which affords a dynamic, mid-sagittal view of the entire vocal tract from lips to larynx.

These results suggest that it is plausible to analyze the the /-Vd/ allomorph as involving same suffix subgraph as for the other two allomorphs, but differing only in how the suffix subgraph is coordinated to the stem. How does the coupling differ these allomorphs? Fig. 10 shows the coupling graph and resulting spectrogram for two coupling alternatives. On the left (a), the coupling is the same as that specified for the other allomorphs— the *TT-clo* gesture of the suffix is shingled to the final constriction gesture of the stem, using eccentric CLO-CLO coupling. The result of this is just a longer /d/ closure (partly voiced and partly voiceless). On the right (b), the*TT-clo* gesture of the suffix is shingled to the final *release* gesture of the stem instead of the closure (REL-CLO coupling). As the spectrogram shows, this coordination yields a (targetless) vowel between the two tongue tip closures, and is therefore a possible model for this allomorph. If this is right, then all the regular past tense forms can be predicted from a single coupling subgraph for the affix, and two qualitatively different coordination patterns—(a) shingling the *TT-clo* of the affix with the final constriction gesture of the stem and (b) shingling the *TT-clo* of the affix with the final release gesture of the stem. Accounting for this qualitative shift in coordination pattern thus becomes the challenge for a dynamical account of the past tense.

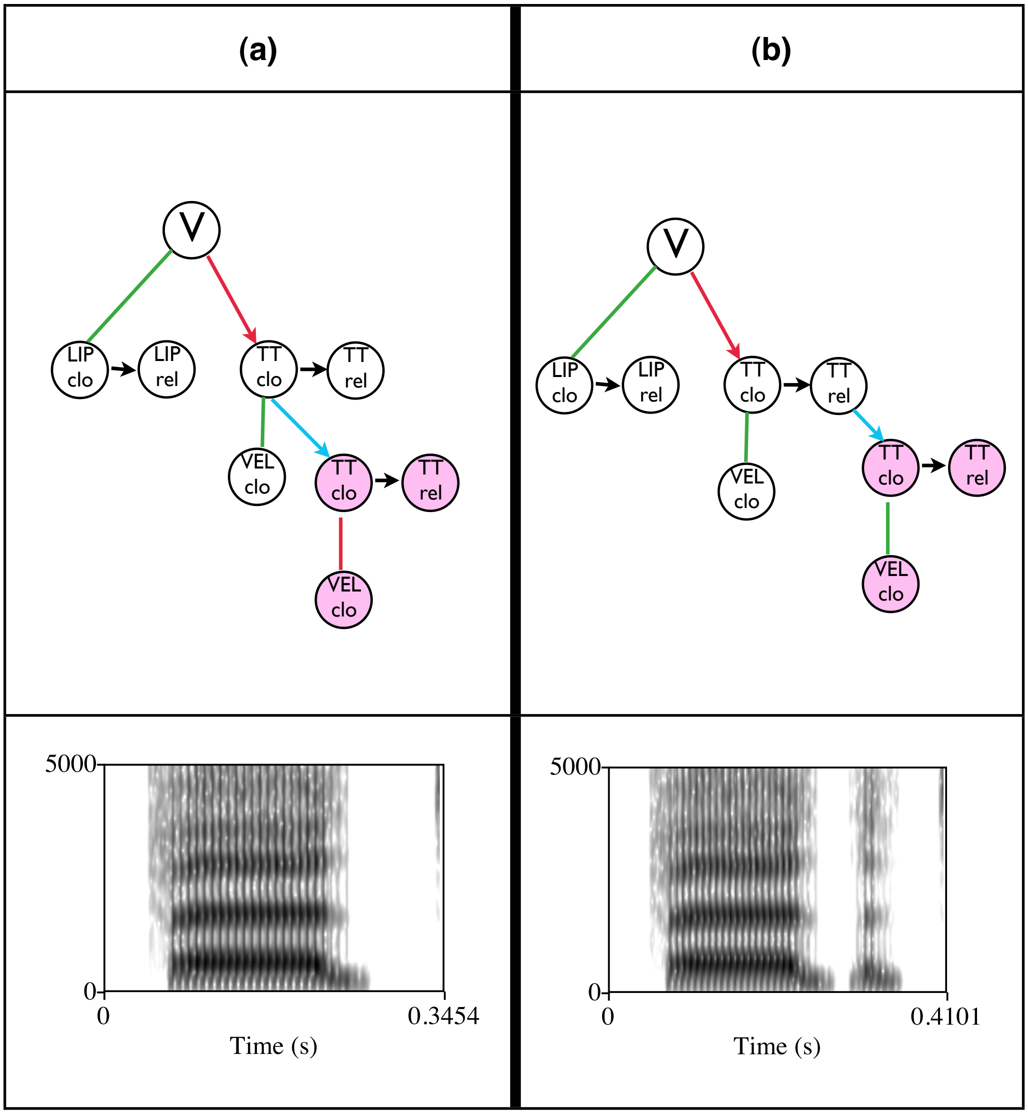

Figure 10. Coupling graph and spectrograms of TaDA output for two models of "padded": (a) CLO-CLO shingling. (b) REL-CLO shingling

IV. Dynamical coordination grammar for past-tense allomorphy

To begin approaching such an account, consider first the outcome of the general coda principle (1) in the case where the stem ends in a TT closure, as in Fig. 10(a). Note that there is no release of the stem's tongue tip gesture visible in the spectrogram, despite the fact when the stem ends in a Lip closure ('nabbed'), as in Fig. 8, the same coupling graph does result in a release of the lips before the TT closure. Why do the two graphs have outputs that differ in this way?

The shingled coordination we have been assuming produces a pattern of gestural activation shown in Fig. 11. As explained earlier, the C2 gesture is activated at a time halfway between the time of activation of C1 and the time it gets to its target. This means that C2 will only get halfway to its target when active control of C1 closure ends, and the release of that constriction begins. So if C1 and C2 are controlling distinct constricting devices (as in the Lip-TT case), the articulatory release of C1's constriction will result in an aerodynamic release of the pressure trapped behind that constriction, and thus a measurable acoustic burst, as we see in Fig. 8. However, in the case where the two gestures control the same constricting device (as in the TT-TT case), at the end of activation of C1 closure, C2 is already active with the goal of producing the same closure as C1, so release of the constriction does not happen, and the result is no burst, as we see in Fig. 10a. (The parameter blending model in TaDA (Saltzman & Munhall, 1989) has the effect of suppressing a release gesture if it is active concurrently with a constriction gesture of the same constricting device). These different consequences of the same C-C coordination pattern were shown by Gafos (2002) to have a variety phonological consequences on the process of template satisfaction in Moroccan Colloquial Arabic (MCA). In the template that requires a CC sequence at the end of a form in a coordination relation similar to that in (1), homorganic C-C sequences are avoided by a variety of phonological mechanisms.

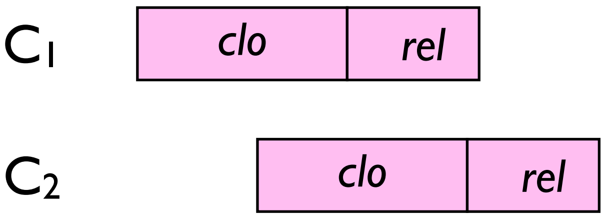

Figure 11. Gestural score for shingled closures

So how can the grammar account for the avoidance of (final) C-C sequences in both MCA and English? Gafos proposes a gestural version of the OCP, banning identical overlapping gestures. However, in English, final homorganic nasal-stop sequences appear to allow the CLO-CLO coordination pattern. Fig. 12 shows this coupling graph and output spectrogram for the word "panned." The coupling is exactly the same as in the non-occurring version of "padded" in Fig. 10a. This graph produces the right result for "panned", while when REL-CLO coupling is employed (not shown), an additional uncontrolled vowel appears in the output. Thus, the relevant OT constraint cannot be stated at the level of gestures. One alternative would be to state that overlapping TT gestures are banned when both are coupled to velic closure gestures. However, this approach appears to miss a generalization. When we consider the behavior of the plural suffix in English, it appears to be quite similar to that of the past tense, in that the plural allomorphs (/z/, /s/, /ɨz/) differ in voicing and syllabicity. The suffix could once again be represented with an invariant gestural sub-graph, with the allomorphy resulting from aerodyamic effects and from a shift (from CLO-CLO to REL-CLO) in how the suffix is coupled to the stem. It would be desirable, therefore, to have a single account for both. However, the environment that triggers the alternate (REL-CLO) coordination pattern in the case of the plural is not limited to identical gestural constellations (e.g., busses), but also includes cases in which the stem ends in an a palatoalveolar sibilant (e.g., "bushes").

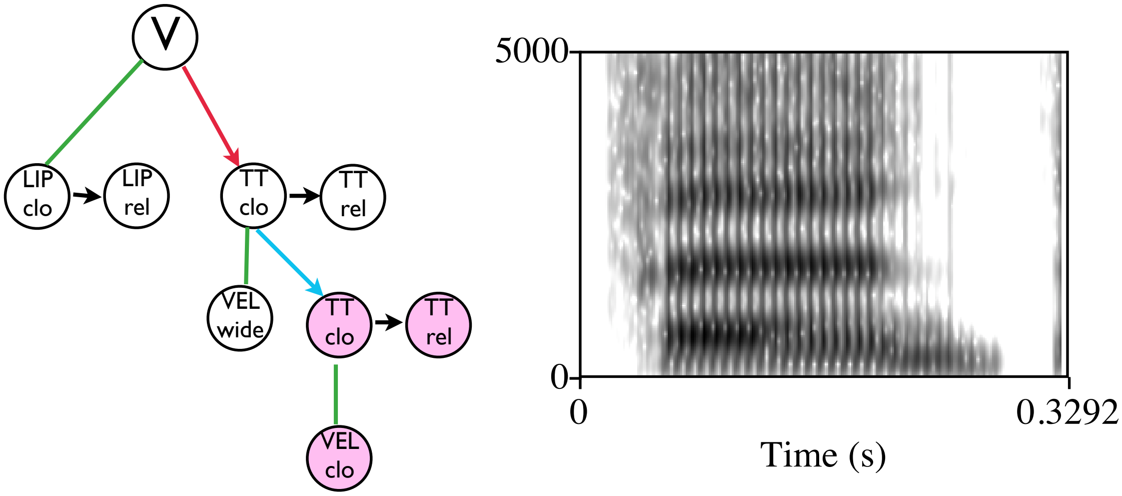

Figure 12. Coupling graph and spectrogram of TaDA output for “panned”

One alternative is to hypothesize that the inflectional suffixation requires some change in the vocal tract output. Evidence that such a principle is operative in English also comes from studies (Guy, 1991; Walker, 2008) showing that final coronal stop deletion in clusters in English is partly conditioned by morphological status: /t/ in a monomorphemic form (e.g. “perfect”) deletes more readily than when it is the past suffix (“packed”). (These apparent deletions may also result from changes in gestural overlap, rather than elimination of the coronal gesture, see Browman & Goldstein, 1990c). A complete formal dynamical account along these lines has not been constructed. Apart from anything else (like the fact that it is a difficult problem), there is insufficient data to do so. Nonetheless, the broad outlines of such an approach will be illustrated here.

Browman & Goldstein (1989) developed the idea that from an aerodynamic point of view, the vocal tract can be modeled abstractly as a set of pipes connected either in series or parallel in a hierarchical arrangement. They showed that Constriction Degree (CD) at various tube levels in the hierarchy— Gesture, Tongue, Oral Cavity, Vocal Tract —may all be relevant to the foundation for phonological natural classes and generalizations. For these purposes, CD is coarsely categorized aerodynamically— Closed, Critical (Turbulence), Open. Using this approach, the past-tense affixation can be associated with the preferred aerodyamics in (2).

(2) Change CD at the VT level in going from stem to affix.

For example in “sewed”, CD changes from Open (for the vowel) to Closed during the coronal closure. In “napped,” there is an Open interval when the lip closure is released before the coronal closure. Of course, it is possible that at some speaking rates, this release disappears. It is exactly this type of data that would be relevant to developing a formal account. Of particular interest would be to see if the probability of release is influenced not only by rate, but also morphological status (e.g., “napped” vs. “apt”).

With a slight sharpening of CD categories, replacing Critical with two categories, corresponding to Wake Turbulence and Channel Turbulence, respectively, the same principle will work with English plurals. The plural “bushes” requires the REL-CLO coupling, because without it there would be no CD change: the end of the stem and the suffix both exhibit wake turbulence. However, “cuffs”, appears to employ CLO-CLO coupling, in which there is a change from channel turbulence to wake turbulence.

In order to instantiate the preference in (2) within the coupled oscillator model, it would be necesary to develop a graph-dynamical model that changes the topology of the coupling graph based on feedback from an internal model of the speech output. Such a model has not yet been developed. However, the grammatical consequences of such a model can be formalized macroscopically using the grammar dynamics approach to constraint interaction (Gafos, 2006; Gafos & Benus, 2006).

In this approach, nonlinear differential equations are used to model the likelihood that the phonological system will settle to one or more *attractor* states along a linguistically-relevant continuous variable, referred to generally as an order parameter. In our case, the order parameter should describe the coordination of the suffix with the stem. To treat this as a (potentially) continuous variable, we allow the suffix *TT-CLO* to be coupled to both the preceding and *CLO* and *REL,* with differing coupling strengths. The order parameter can then be defined as their *relative* coupling strength (defined as the log of the ratio: strength(REL-CLO) / strength(CLO-CLO)). Thus, CLO-CLO coupling would have a negative value of the order parameter, while CLO-REL would have a positive value.

The differential equation required is one with at least two possible attractor states. A cubic polynomial is the simplest function that can accomplish this. The tilted anharmonic oscillator is one such function, and is shown (slightly simplified form) in (3):

(3) $$\dot{x} = f(x, R) = R + x - x^3$$

where, represents x the value of the order parameter, $\dot{x}$ is its time-derivative, and R is a constant that influences the layout of attractor states in a manner to be shown below. This system has been used to model perceptual categorization of speech (Tuller et al., 1994), vowel harmony in Hungarian (Gafos & Benus, 2006), and the location of nuclear pitch accent in English vs. Spanish (Nava, 2010). The stable states of a such system can be graphically examined by plotting the *potential* of the system $V(x)$ as a function of the value of the order parameter. Since the change in the order parameter is proportional to the slope of the potential, values of the order parameter where the potential has zero slope are stable states. The potential function ($V(x)$) associated with $f(x,R)$ is plotted in Fig. 13, for the case where R = 0.

The way to interpret these figures is that the potential function defines a surface, and the order parameter behaves like a ball dropped onto that surface. There are two stable states where the ball can settle, one with a positive value of x (corresponding to *REL-CLO* coupling), and one with a negative value of x (corresponding to *CLO-CLO* coupling). Of course, these two stable states are not equivalent; they occur in different contexts. This contextual effect can be modeled by the value of R, which is a control parameter that governs the overall tilt of the potential.

Setting R<0 results in a single stable state at a negative value of relative coupling strength (CLO-CLO coupling), while setting R>0 results in a single stable state with a positive value of relative coupling strength (REL-CLO coupling). Fig. 14 shows the effect of setting $R=-1$ on the left, and $R=1$ on the right. Thus the *quantitative* value of R acts like a knob that can be used to switch the system between its two stable states. The value of R could be set to a negative value in the general case, but it is turned up to some positive value when the output would not yield a change in CD.

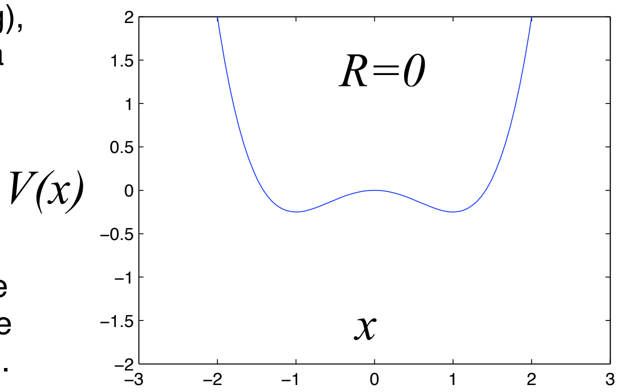

Figure 13. Potential function of Equation (3) when R=0

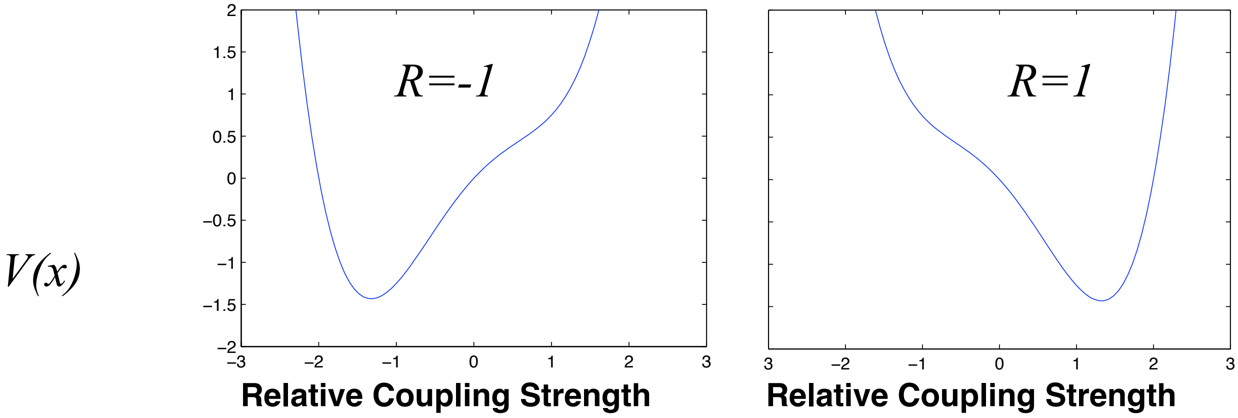

Figure 14. Potential function for the log of relative coupling strength of CLO-CLO and REL-CLO coupling

This analysis may appear to have a lot of theoretical baggage, just to switch coordination patterns between two different contexts. Why is this preferable to a learned rule that sets the coordination appropriately in the defined contexts? The advantage of the dynamical model would be demonstrated if the selection of coupling modes depended on some *quantitative* variable, as well as the qualitative morpho-phonological conditioning of dependence on CD change. To see a possible role for quantitative conditioning, we need to consider some irregular past-tense forms, for words like those in (4):

(4) shed, spread, put, let, set, cut, hit, beat, shut, hurt, cost, cast, burst

While these are usually thought of as having an irregular zero-suffix, it is possible to analyze them as having the regular past-tense subgraph, but an irregular coordination pattern, i.e, they choose the CLO-CLO coordination pattern, rather than the REL-CLO pattern. (We ignore for the present the expectation that if they are formed with a suffix and the CLO-CLO coordination, the past-tense final stop should be a bit longer than that of the bare stem, another point on which there is no data) In addtion, there are verbs that (in my experience) can be produced either using CLO-CLO (irregularly) or REL-CLO, for example, “slot” (I have heard both “slot” and “slotted”) and possibly “bat” (in a baseball context--”She bat .300 last year”). Also in Albright and Hayes’ (2003) wug-test study of English past tense, participants gave variable responses to the nonce verbs in (5) that included a substantial (>10%) number of responses in which there was no change in the past-tense form, which can be analyzed as resulting from CLO-CLO coupling.

(5) gleed, glit, gude

Suppose an experiment like Albright and Hayes’ were undertaken under conditions that manipulated speaking rate. It would not be surprising if speaking rate pushed the distribution of responses in the direction of the shorter forms that result from CLO-CLO coordination. This speaking rate effect could then be modeled by making the value of R dependent on speaking rate as well as CD change.

V. Conclusion

We have presented an analysis of the regular English past tense allomorphy, in which the suffix is represented by one invariant coupling graph that couples to the stem in two qualitatively distinct coupling links. A nonlinear grammar-dynamical system that can switch the coupling modes was sketched, though its details are missing. It is striking that pursuing an explicit dynamical account of something as apparently simple as the regular English past tense revealed not only some small new insights and generalizations, but also revealed just how much we do not know and what kinds of new experiments might inform us further.

* This work was supported by NIH NIDCD DC008780. Many thanks to Khalil Iskarous, Adam Lammert, Jean-Roger Vergnaud and Hosung Nam for helpful discussion, and to Emily Nava for comments on an earlier draft. The title of the paper has two personal senses (in addition to a third which is a pop culture reference to movies such as “Back to the Future.”) One sense is a re-connection with Judy Aissen, along the lines we first connected. A second sense is that a very preliminary version of this paper was drafted by Cathe Browman and myself, for a talk she presented at the International Congress of Linguists in Prague in 1992 (I believe). The text has been lost in the sands (or rather electrons, I suppose) of time. When we wrote it, we did not have much data, and the model had not developed to the point that we could evaluate acoustic output. But the basic ideas were pretty much there. So this is really her paper, as much as it is mine.

References

Albright, A. & Hayes, B. (2003). Rules vs. analogy in English past tenses: a computational/experimental study. *Cognition, 90,* 119–161.

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

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

Browman, C. P. & Goldstein, L. (1990b). Representation and reality: physical systems and phonological structure. *Journal of Phonetics, 18,* 411-424.

Browman, C. P., and Goldstein, L. (1990c). Tiers in articulatory phonology, with some inplications for casual speech. In J. Kingston and M. E. Beckman (eds.) *Papers in Laboratory Phonology, vol. 1, Between the Grammar and Physics of Speech.* Cambridge, UK: Cambridge University Press, pp. 341–376.

Browman, C. P., and Goldstein, L. (1992a). Articulatory phonology: an overview. *Phonetica, 49*, 155–180.

Browman, C. P. & Goldstein, L. (1992b). "Targetless" schwa: an articulatory analysis. In Docherty, G. J. and D. R. Ladd, *Papers in Laboratory Phonology II: Gesture, segment, prosody* . Cambridge, UK: Cambridge University Press, pp. 26-56.

Browman, C. P., and Goldstein, L. (1995). Dynamics and articulatory phonology. In T. van Gelder (ed.) *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press, pp. 175–194.

Browman, C.P., & Goldstein, L. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée, 5,* 25–34.

Byrd, D. (1995). C-centers revisited. *Phonetica*, 52, 285–306.

Flemming, E. & Johnson, S. (2007). Rosa'sroses: reduced vowels in American English, *Journal of the International Phonetic Association, 37*, 83-96.

Gafos, A. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory, 20*, 269-337.

Gafos, A. (2006). *Dynamics in Grammar: Comment on Ladd and Ernestus & Baayen. Laboratory Phonology 8: Varieties of Phonological Competence,* In Goldstein, L., Whalen, D. H. and Best, C. (Eds.) Berlin/New York: Mouton de Gruyter, pp. 51-79.

Gafos, A. & S. Benus. (2006). The dynamics of phonological cognition. *Cognitive Science* 30, 5, 905-943.

Gafos, A. & Goldstein, L. (to appear). Articulatory representation and organization. In Cohn, A., Fourgeron, C., & Huffman, M. (Eds.) *Handbook of Laboratory Phonology.*

Gafos, A. & Kirov, C. (2009). A dynamical model of change in phonological representations: The case of lenition. In F. Pellegrino, E. Marisco, & I. Chitoran, (Eds). *Approaches to phonological complexity.* Berlin/New York: Mouton de Gruyter. pp. 219-240.

Goldstein, L., Byrd, D., and Saltzman, E. (2006) The role of vocal tract gestural action units in understanding the evolution of phonology. In M. Arbib (Ed.) *From Action to Language: The Mirror Neuron System.* Cambridge: Cambridge University Press. pp. 215-249.

Goldstein, L., Chitoran, I, & Selkirk, E. (2007). Syllable structure as coupled oscillator modes: Evidence from Georgian and Tashlhiyt Berber. Trouvain, W. and Barry, W. (Eds). *Proceedings of the XVI International Congress of Phonetic Sciences,* pp. 241-244.

Goldstein, L., Nam, H., Saltzman, E., & Chitoran, I. (2009). Coupled Oscillator Planning Model of Speech Timing and Syllable Structure. In G. Fant, H. Fujisaki, & J. Shen, (Eds.). *Frontiers in phonetics and speech science*, Beijing: The Commercial Press. Pp. 239-250.

Goldstein, L, Pouplier, M., Chen, L., Saltzman, E., and Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition, 103*, 386-412.

Guy, G.R. (1991). Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change, 3,* 1-22.

Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51*, 347-356.

Hanson, H.M. & Stevens, K.N. (2002). A quasiarticulatory approach to controlling acoustic source parameters in a Klatt-type formant synthesizer using HLsyn. *J Acoust Soc Am. 112,* 1158-82.

Hermes, A., Grice, M., Mücke, D., & Niemann, H. (2008). Articulatory indicators of syllable affiliation in word initial consonant clusters in Italian. In R. Sock, S. Fuchs & Y. Laprie (Eds.), Proceedings of the 8th International Seminar on Speech Production. (pp. 433-436). Strasbourg.

Honorof, D., & Browman, C.P. (1995). The center or edge: How are consonant clusters organized with respect to the vowel? In K. Elenius & P. Branderud (Eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm, Sweden (pp. 552-555). Stockholm: KTH and Stockholm University.

Hsieh, F.-Y. (2010). Gesture Reorganization in Mandarin Tone 3 Sandhi. Presented at the meeting of the Acoustical Society of America, Cancun.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior.* Cambridge, MA: MIT Press.

Kelso, J. A. S., Scholz, J. P., & Schöner, G. (1986). Nonequilibrium phase transitions in coordinated biological motion: Critical fluctuations. *Physics Letters, 118*, 279-284.

Kugler, E N., & Turvey, M. T. (1987). I*nformation, natural law, and the self-assembly of rhythmic movement*. Hillsdale, NJ: Erlbaum.

Lammert, A., Goldstein, L. & Narayanan, S. (2010). Gestural control in the English past-tense suffix: an articulatory study using real-time MRI. Presented at 12th Conference on Laboratory Phonology. Albuquerque, NM/

Löfqvist, A., Gracco, V. 1999. Interarticulator programming in VCV sequences: lip and tongue movements. Journal of the Acoustic Society of America 105, 1854-1876.

Marin, S, & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control, 14*, 380-407.

Nam, H., (2007). Syllable-level intergestural timing model: Split-gesture dynamics focusing on positional asymmetry and moraic structure. In: Cole, J., Hualde, I. J. (Eds.), *Laboratory Phonology 9*. Mouton de Gruyter, New York, pp. 483–506.

Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: a coupled oscillator model. In F. Pellegrino, E. Marisco, & I. Chitoran, (Eds). *Approaches to phonological complexity.* Berlin/New York: Mouton de Gruyter. pp. 299-328.

Nam, H., Goldstein, L., Saltzman, E., and Byrd, D. (2004). TADA: An enhanced, portable Task Dynamics model in MATLAB. *J. Acoust. Soc. Am. 115*, 2430 (abstract).

Nava, E. (2010). Connecting Phrasal and Rhythmic Events: Evidence from Second Language Speech. PhD Dissertation. University of Southern California.

Parrell, B. (submitted). The role of gestural phasing in Western Andalusian Spanish aspiration. *Journal of Phonetics.*

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

Saltzman, E., Nam, H., Krivokapic, J., & Goldstein, L. (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In P. A. Barbosa, S. Madureira, & C. Reis,Â (Eds.), *Proceedings of the 4th International Conference on Speech Prosody (Speech Prosody 2008),* Campinas, Brazil.

Shaw, J., Gafos, A., Hoole, P., & Zeroual, C. (2009). Syllabification in Moroccan Arabic : evidence from patterns of temporal stability in articulation. *Phonology, 26*, 187–215.

Smorodinsky, I. (2001). Schwas with and without active gestural control. *Journal of the Acoustical Society of America, 109*, 2446 (abstract).

Torreira, F. (2007a). Pre- and postaspirated stops in Andalusian Spanish . In P. Prieto, J. Mascaró, & M.-J. Solé (Ed.), *Segmental and prosodic issues in Romance phonology* (pp. 67-82). Amsterdam/Philadelphia: John Benjamins.

Torreira, F. (2007b). Coarticulation between aspirated-s and voiceless stops in Spanish: an interdialectal comparison. In N. Sagarra, & A. J. Toribio (Eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium* (pp. 113-120). Somerville, MA: Cascadilla Press.

Trager, G. L. & Smith, H. L. (1951). *An Outline of English Structure (Studies in Linguistics: Occasional Papers 3)*. Norman, OK: Battenburg Press.

Trubetzkoy, N. (1939). Grundzüge der Phonologie. *Travaux du cercle linguistique de Prague 7.*

Tuller, B., Case, P., Ding, M., & Kelso, J. A. S. (1994). The nonlinear dynamics of speech categorization. *Journal of Experimental Psychology, 20*, 3–16.

Turvey, M., 1990. Coordination. *Am. Psychologist 45:* 938–953.

Walker, J.A. (2008). Form, function, and frequency in phonology: (t/d)-Deletion in Toronto. *NWAV, 37,* Houston, TX.

Yanagawa, M.(2006). *Articulatory Timing in First and Second Language: A Cross-Linguistic Study*. PhD Dissertation, Yale University.

Zsiga, E.C. (1995). An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English," in B. Connell and A. Arvaniti (eds.), *Phonology and Phonetic Evidence:Papers in Laboratory Phonology IV,* Cambridge: Cambridge University Press, pp. 282 - 302.

Zsiga, E.C. (2000). Phonetic alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics*, 28 , 69-102.