1200

ELSEVIER

Human Movement Science 19 (2000) 499–526

HUMAN MOVEMENT SCIENCE

www.elsevier.com/locate/humov

Task-dynamics of gestural timing: Phase windows and multifrequency rhythms

Elliot Saltzman [a,c,*], Dani Byrd [b,c]

[a] *Department of Physical Therapy, Boston University, 635 Commonwealth Avenue, Boston, MA 022115, USA*

[b] *Department of Linguistics, University of Southern California, Los Angeles, CA, USA*

[c] *Haskins Laboratories, New Haven, CT, USA*

Abstract

In this paper, we explore the hypothesis that intergestural phasing relations are implemented via coupling terms in a nonlinear dynamical systems model. Specifically, we describe recent computational developments of the *task-dynamic* model of gestural patterning (e.g., E. Saltzman, J. A. S. Kelso, Psychological Review 94 (1987) 84–106; E. Saltzman, K. G. Munhall, Ecological Psychology 1 (1989) 333–382) that are focused on modeling the timing of rhythmic action units. First, we explore the possibility of attractor states for intergestural phasing that are characterized as ranges or *phase windows* (D. Byrd, Phonology 13 (1996) 139–169), and contrast this behavior with standard models that display punctate relative phasing. It is argued that the phase window approach can provide flexible control of the relative timing of articulatory gestures, allowing constrained variability in intergestural timing as a function of linguistic and para-linguistic factors. Second, we discuss how this extension of the task-dynamic model has been adapted for modeling the production of multifrequency rhythms (speech or bimanual). This work explores the control of the frequency- and phase-locking characteristics of coupled limit cycle oscillators by examining how desired frequency ratio, intrinsic frequency detuning, and coupling asymmetries interact in creating observed rhythmic patterns. Using this method, details of the resultant transient and steady-state trajectories of phase and amplitude are generated that are not available

* Corresponding author. Tel. +1-617-353-7494; fax: +1-617-353-9463.

PII: S0167-9457(00)00030-0

using models derived with averaging techniques.

PsycINFO classification: 2330; 2720

Keywords: Multifrequency rhythms; Coupled oscillators; Task dynamics; Phase windows; Speech timing; Relative phasing

1. Introduction

The temporal patterns of speech production can be described according to four types of timing properties: intragestural, transgestural, global, and intergestural. *Intragestural* timing refers to the temporal properties of a given gesture [1], e.g., the time from gestural onset to peak velocity (e.g., Byrd & Saltzman, 1998; Bullock & Grossberg, 1988; Coker, 1976; Kröger, Schröder & Opgen-Rhein, 1995; Ostry, Gribble & Gracco, 1996) or to target attainment (Abry & Lallouache, 1995; Bell-Berti & Harris, 1981; Henke, 1966; Lubker, 1981; Perkell & Matthies, 1992; Saltzman, 1995; Saltzman & Mitra, 1998; Rubin et al., 1996; Saltzman, Löfqvist & Mitra, 2000). *Transgestural* timing refers to modulations of the timing properties of all gestures active during a relatively localized portion of an utterance, and has been hypothesized to result from transient modulation of the stiffness parameters of all currently active gestures (e.g., Byrd, Kaun, Narayanan & Saltzman, 2000) or of the timeflow speed in an hypothesized underlying central "clock" that governs the ongoing rate or style of speaking (e.g., Port, 1986; Port & Cummins, 1992; Saltzman, Löfqvist, Kay, Kinsella-Shaw & Rubin, 1998; Saltzman et al., 2000; see also Nguyen & Cottrell, 1997). The settings of the clock's timeflow speed are hypothesized to underlie *global* timing properties, i.e., overall speaking rate or style, at the timescale of an entire utterance. The fourth timing property – *intergestural* – refers to the patterns of relative timing or phasing among an utterance's component gestures.

These four timing properties are inter-related. For example, transgestural and global timing effects will alter the intragestural timing properties of the affected gestures. Similarly, transient increases in clock speed induced, for

[1] The term gesture denotes a member of an equivalence class of articulatory movement patterns that are controlled with reference to a given task-relevant goal (e.g., a bilabial closure during speech (Saltzman & Munhall, 1989), or striking a piano key during musical performance).

example, by mechanical perturbations of the articulators, do not only affect intragestural timing but also result in large transient changes, and smaller (by an order of magnitude) steady-state changes, of intergestural timing patterns (Saltzman et al., 1998). Finally, discontinuous transitions of intergestural phasing have been demonstrated as global speaking rate is increased beyond certain critical values (Kelso, Saltzman & Tuller, 1986a,b; Tuller & Kelso, 1991).

In this paper, we first explore the manner in which intergestural phasing relations are implemented using a pair of 1:1 frequency-locked oscillators via coupling terms in a nonlinear dynamical systems model. Specifically, we describe recent computational developments of the *task-dynamic* model of gestural patterning (e.g., Saltzman & Kelso, 1987; Saltzman & Munhall, 1989) that are focused on modeling the intergestural timing properties of rhythmic action units, and describe the manner in which the relative phasing of speech gestures is related conceptually to the intergestural phasing patterns displayed by coupled rhythmic action units. We then describe how the model has been extended to include the possibility of attractor states for intergestural phasing during speech that are characterized by *phase windows* (Byrd, 1996a), and contrast this behavior with standard models that display punctate relative phasing. Finally, we discuss how the task-dynamic control of intergestural phasing has been generalized to the production of multifrequency (*m:n* frequency-locked) rhythms.

2. Task dynamics and phase windows

In the task-dynamic model of speech production (Saltzman & Munhall, 1989), articulatory movement patterns are conceived of as coordinated, goal-directed gestures that are dynamically defined. In particular, they have been modeled as critically damped oscillators that act as point attractors. An utterance is described not only by the parameter values of the component gestures but by how the gestures are coordinated or phased with one other. This approach captures coproduction by allowing gestures to overlap in time, with the acoustic consequences of the coproduced units reflecting their combined influence on the vocal tract.

In this model, the spatiotemporal patterns of speech emerge as behaviors implicit in a dynamical system with two functionally distinct but interacting levels (Fig. 1). The *interarticulator* level is defined according to both *model articulator* (e.g. lips and jaw) coordinates and *tract-variable* (e.g. lip aperture

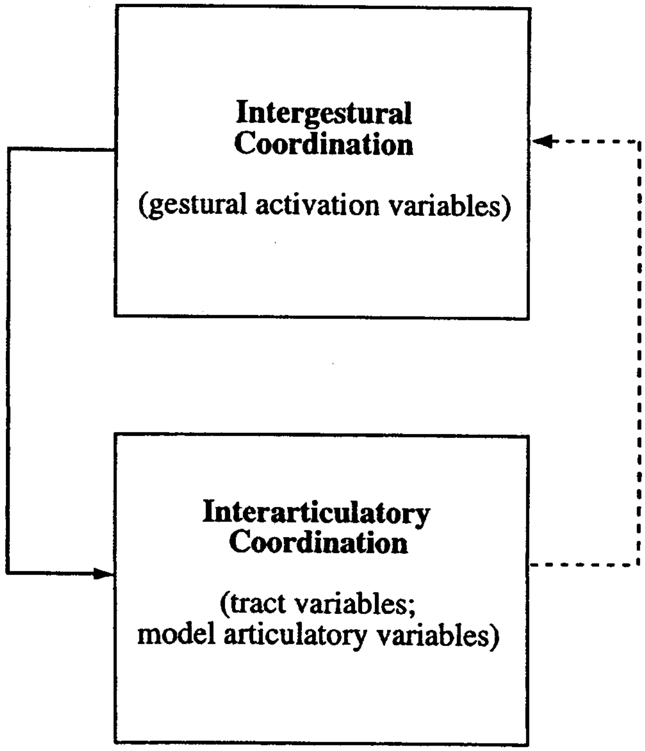

Fig. 1. Schematic illustration of the proposed two-level dynamical model for speech production, with associated coordinate systems indicated. The solid arrow from the intergestural to the interarticulator level denotes the feed forward flow of gestural activation. The dashed arrow indicates feedback of ongoing tract-variable and model articulator state information to the intergestural level.

and protrusion) coordinates. The *intergestural* level is defined according to a set of *activation* coordinates. Invariant gestural units are posited in the form of context-independent sets of dynamical parameters (e.g. target, stiffness, and damping coefficients) and are associated with corresponding subsets of model articulator, tract-variable, and activation coordinates. Each unit's activation coordinate reflects the strength with which the associated gesture "attempts" to shape vocal tract movements at any given point in time. The tract-variable and model articulator coordinates of each unit specify, respectively, the particular vocal-tract constriction (e.g. bilabial) and articulatory synergy (e.g. lips and jaw) whose behaviors are affected directly by the associated unit's activation. The interarticulator level accounts for the coordination among articulators at a given point in time due to the currently active gesture set. The intergestural level governs the patterns of relative timing among the gestural units participating in an utterance and the temporal evolution of the activation trajectories of individual gestures in the utterance.

Phase has been adopted as a measure of intrinsic gestural time in speech gestures despite the fact that point attractor models have been used to model these gestures. The coordination of gestures in terms of their relative phase has to date been specified "by hand" in models of word production as part of an utterance's *gestural score* (Browman & Goldstein, 1990). Formal implementations of relative phasing among articulatory gestures have to date been in terms of invariant point-to-point, or *punctate*, phase relations such as aligning constriction onset to target or constriction onset to release. Ultimately a model of speech timing must be developed that allows relative phasing to self-organize as it does in coupled oscillatory systems. As we discuss below, relative timing in speech must allow constrained variability. But the best developed theories of inter-unit timing come from work in (nonspeech) rhythmic movement and have been limited to situations in which the movements display specific point-to-point relative phase relations. We assume that rhythmic and non-rhythmic speech behaviors have a common underlying dynamical organization. Here we attempt to reconcile work in coupled oscillator dynamics and intergestural timing in speech.

Coordination of gestures in speech refers to both timing within a traditional speech segment and timing between speech segments. Within a segment, for example, oral closure and laryngeal adduction in the case of stop consonants, or lip rounding and tongue rear constriction for a [w], are critically timed. Gestures for sequential speech sounds must also be coordinated both within and between words. These might include sequential consonant constrictions in a consonant cluster such as [sk] or vowels and consonants in sequence. It has been observed that the relative phasing among articulatory gestures is not fixed, i.e., the target phasing values are not single invariant points such as 180°, 229°, 240°. Rather, the observed values of intergestural relative phase are affected by linguistic factors such as prosody, stress, syllabification, etc. Studies of intergestural coordination in speech have repeatedly demonstrated systematic variability as a function of linguistic and extra-linguistic factors (e.g. Lubker, 1986; Nittrouer, 1991; Nittrouer, Munhall, Kelso, Tuller & Harris, 1988), such as syllable structure (Browman & Goldstein, 1990; Byrd, 1996b; Krakow, 1993; Sproat & Fujimura, 1993), stress (Nittrouer et al., 1988; Shaiman & Porter, 1991), segment identity (e.g. Browman & Goldstein, 1990; Byrd, 1996b), phrasal structure (e.g. Byrd et al., 2000; Hardcastle, 1985), and speaking rate (e.g. Byrd & Tan, 1996; Shaiman, Adams & Kimelman, 1995).

Given this systematic, context-specific variation of intergestural phasing, Byrd (1996a) has hypothesized that the target values are not "punctate" but are constrained instead to fall within an admissible range or *phase window*. For example, if the system's phase window is defined from 220° to 245°, its steady-state relative phase values can fall anywhere within this window, depending on contextual factors such as the initial distance from the window, the intrinsic gestural frequencies, and the "attractor strength" of the intergestural dynamics. The reasons for this constrained and principled variability may lie in the intrinsic communicative goals of speech, given that articulatory movement patterns are governed in part by the perceptual capacities of listeners and the competing demands which the speaker-listener relationship necessitates. That is, temporal relations among phonological units are one means of communicating to the listener both the segmental identity of those units and their position in the hierarchy of linguistic syllabic and phrasal structure.

When a task dynamics for speech motor control was first proposed, Saltzman & Munhall (1989, p. 365) hypothesized that "gestures cohere in bundles corresponding, roughly, to traditional segmental descriptions, and that these segmental units maintain their integrity in fluent speech." They attributed this cohesion to dynamical coupling of the gestures. Such cohesion is important – for example, if the timing between the tongue dorsum gesture and the primary place gesture in a click consonant is not strictly controlled, that segment will not be produced adequately, namely, with the requisite velaric ingressive airstream. (See Byrd, 1996a, for other examples of requisite intra-segmental timing.) No such dramatic constraints exist between gestures of separate segments or words, and evidence suggests that gestures forming a segment stand in a substantially tighter relationship than those which participate in different segments (Munhall, Löfqvist & Kelso, 1994; Saltzman et al., 2000). In the present work we demonstrate that an appropriate type of coupling can yield a window of relative phasing between two coupled oscillators, lending computational plausibility to the hypothesis that gestures belonging to single segments are coupled so as to yield narrow phase windows, while those belonging to different segments are coupled to yield wider phase windows (Byrd, 1996a). By extension, cross-word timing may be even more variable, i.e., controlled by even wider phase windows.

In the following section, we explore the hypothesis that intergestural phasing relations are implemented via coupling terms in a nonlinear dynamical systems model. In a set of simulations, we contrast the behavior of a standard model of coupled oscillators (limit-cycles) that displays punctate relative phasing with an extended model of relative phasing that can display either stable phase points or windows.

3. Modeling relative phasing: Phase points and windows

In this section, we present a method of flexibly controlling 1:1 frequency-locking and relative phasing among a pair of coupled oscillators in a manner that can produce stable point or window intergestural phasing patterns in the steady-state. This approach generalizes previous mathematical treatments in the motor control literature that use nonlinear dynamical systems theory and are restricted to the production of punctate relative phasing patterns. It also represents an improvement over the method described by Saltzman & Byrd (1999) which was not generalizable to the control of multifrequency rhythms.

We begin by defining the relative phase, ψ, between two oscillators as an abstract *task-space* variable (or *order parameter*), while the ongoing motion of each component oscillator is defined in terms of a more physically-based set of *model articulator* variables. Relative phase, the state variable at the task-space level, is defined as

$$\psi = \phi_2 - \phi_1, \tag{1}$$

where ϕ_i is the phase angle of the ith oscillator. The articulator variables for each oscillator are simply the respective sets of (position x, and velocity $\dot{x}$) state variables at the component oscillator level. We adopt a so-called *hybrid* oscillatory dynamics (Kay, Kelso, Saltzman & Schöner, 1987) at the model articulator level, defined by second order equations of motion (see also Haken, Kelso & Bunz, 1985):

$$\ddot{x}_i = -\alpha_i \dot{x}_i - \beta_i x_i^2 \dot{x}_i - \gamma_i \dot{x}_i^3 - \omega_{0i}^2 x_i, \tag{2}$$

where i denotes oscillator identity ($i = 1, 2$), α_i is the linear damping, β_i the nonlinear van der Pol damping, γ_i the nonlinear Rayleigh damping, and ω_{0i} is the linear natural frequency. These equations define independent, *uncoupled* motions for each oscillator's intrinsic dynamics, and are widely used in the movement science literature to model the behavior of coupled rhythmic movements.

In order for the oscillators to exhibit coordinated behavior with a desired pattern of relative phasing, the intrinsic model articulator equations must be coupled to one another in a manner specific to the desired phasing. We do this, roughly, by fashioning a corresponding task-specific dynamics in the relative-phase task-space and transforming these dynamics into an appropriate set of task-specific coupling influences at the model articulator level. We begin by describing our approach in the context of producing punctate patterns of relative phase, in which the coupled system settles to particular relative phase values specified as single points in task space. We then show how our approach readily generalizes to systems that settle to ranges of relative phase values specified as windows in the task space.

3.1. Relative phasing: Target points

Punctate relative phasing patterns can be produced by defining *potential functions*, $V(\psi)$, in the task space with single minima corresponding to the desired relative phase value, ψ_0. One of the simplest such potential functions (Fig. 2, left panel) is

$$V(\psi) = -a\cos(\psi - \psi_0), \tag{3}$$

where the height of the potential valley is equal to *2a*. This potential function can be used to define a first-order equation of motion in task space according to the following general expression:

$$\dot{\psi}_T = -dV/d\psi = -a\sin(\psi - \psi_0), \tag{4}$$

where $\dot{\psi}_T$ is the task-specific relative phase velocity (i.e., the system's desired task-level state velocity) that is a function of the current relative phase (i.e., task-space state).

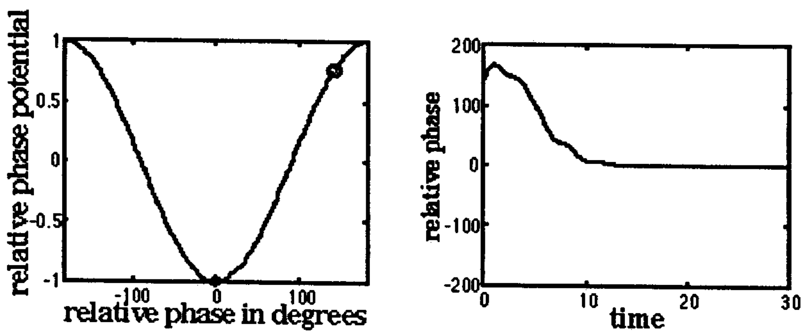

Fig. 2. Left: Task-space potential for punctate relative phasing. Right: Relative phase trajectory. 'o' = initial relative phase, '*' = final relative phase.

The resulting behavior of relative phase can be visualized as that of a ball rolling down the sides of the (sticky) walls of the potential surface until the ball comes to rest at the surface's minimum, ψ_0. We then convert this task-space state velocity into a corresponding set of model articulator coupling terms. This is done in several steps.

The first step is to transform each oscillator's current model articulator state, defined in terms of Cartesian position (x_i) and velocity ($\dot{x}_i$) coordinates on their respective phase planes, into a corresponding set of polar phase angle (ϕ_i) and radial amplitude (A_i) coordinates according to:

$$\phi_i = -\tan^{-1}\left(\frac{\dot{x}_i/\omega_{0i}}{x_i}\right), \tag{5a}$$

$$A_i = \sqrt{x_i^2 + (\dot{x}_i/\omega_{0i})^2}. \tag{5b}$$

The articulator phase angles are used to define the current value for relative phase in the task space using Eq. (1). The desired relative phase velocity, $\dot{\psi}_{\mathrm{T}}$, (i.e., task-specific "forcing") is then defined according to Eq. (4).

The next step is to transform the task-space state velocity, $\dot{\psi}_{\mathrm{T}}$, defined by Eq. (4) into a corresponding set of task-specific coupling terms that contribute to each oscillator's phase velocity. This is done by differentiating Eq. (1) with respect to time and defining the inverse of the resultant expression. Differentiating Eq. (1) yields

$$\dot{\psi} = \dot{\phi}_2 - \dot{\phi}_1 = J\dot{\tilde{\phi}}, \tag{6}$$

where $J = [-1, 1]$, the 1×2 Jacobian matrix of Eq. (1), and $\dot{\tilde{\phi}} = [\dot{\phi}_1, \dot{\phi}_2]^{\mathrm{T}}$, the vector of phase velocities. Inverting Eq. (6) yields

$$\dot{\tilde{\phi}} = J^+\dot{\psi} = [J_1^+, J_2^+]^{\mathrm{T}}\dot{\psi} = [-1, 1]^{\mathrm{T}}\frac{1}{2}\dot{\psi} = [-0.5, 0.5]^{\mathrm{T}}\dot{\psi}, \tag{7}$$

where $J^+ = J^{\mathrm{T}}(JJ^{\mathrm{T}})^{-1}$, the 2×1 unweighted pseudoinverse of the Jacobian, which provides a unique least squares solution for the set of computed component phase velocities used at each point in time.

Combining Eqs. (6) and (7), we get the following expression for the task-specific coupling terms for oscillator-i's phase velocity:

$$\dot{\phi}_{i,T} = J_i^+\dot{\psi}_{\mathrm{T}}. \tag{8}$$

These task-specific phase velocities are transformed into corresponding task-specific coupling terms at the component level:

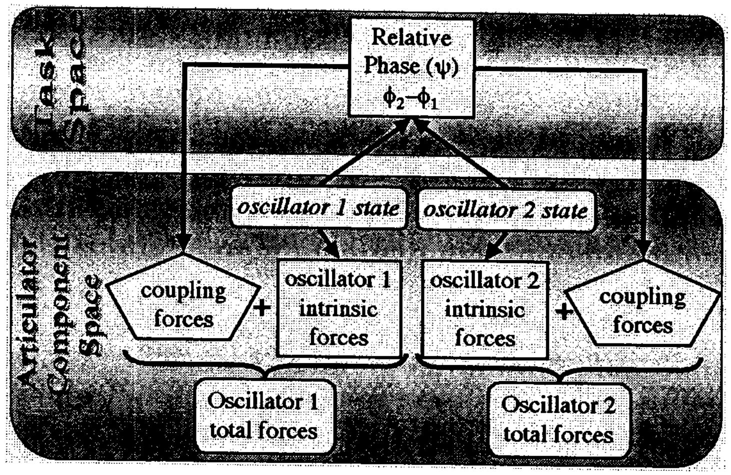

Fig. 3. A graphical schema of the task-dynamics coupled oscillator model.

$$\ddot{x}_{i,T} = \left(\frac{\partial \dot{x}_i}{\partial \phi_i}\right)\dot{\phi}_{i,T}, \quad \text{where} \begin{cases} \dot{x}_i = -\omega_{0i}A_i \sin \phi_i, \\ \left(\frac{\partial \dot{x}_i}{\partial \phi_i}\right) = -\omega_{0i}A_i \cos \phi_i. \end{cases} \tag{9}$$

Combining Eqs. (4), (8) and (9) for the point potential coupling forces yields

$$\ddot{x}_{i,T} = (\omega_{0i}A_i \cos \phi_i)J_i^{+}(a \sin[\psi - \psi_0]). \tag{10}$$

Finally, these task-specific coupling forces are added to the right-hand sides of the intrinsic component equations of motions (Eq. (2)), which are integrated numerically to generate position and velocity trajectories for each component oscillator. All numerical integration was performed using the *ode45* solver (fourth- and fifth-order Runge–Kutta) in Matlab 5.2.1 for the Macintosh. The overall computational method is graphically schematized in Fig. 3 [2].

3.2. Relative phasing: Target windows

In order to generalize the model to display stable phase *windows* in the steady state, we simply use the following task-space potential function (cf., Eq. (3)):

[2] The model structures schematized in Figs. 1 and 3 are related in the following manner. The task level in Fig. 3 corresponds to the intergestural level in Fig. 1 and governs the pattern of relative timing among the component gestural oscillator units through the specification of corresponding interunit coupling functions. It is assumed in the simulations that the oscillator units are always active. The articulator component level in Fig. 3 corresponds to the interarticulator level in Fig. 1.

$$V(\psi) = \frac{2}{1 + e^{-\lambda z}}, \quad \text{where } z = -(\cos[\psi - \psi_0] + 1). \tag{11}$$

This equation represents a cosine-shaped potential whose valley has been flattened by the logistic "squashing function", commonly used in connectionist computational models, in which λ defines the severity of squashing (Fig. 4, left panel; see Large & Kolen (1994) for a related use of squashing functions in a coupled oscillator framework). We chose this function for its relative simplicity. More elaborate functions would allow window width and flatness to be manipulated independently. As with Eq. (3), we use this potential function to define a first-order equation of motion in the task space:

$$\dot{\psi}_{\mathrm{T}} = -\mathrm{d}V/\mathrm{d}\psi = \frac{-2\lambda e^{-\lambda z}}{(1 + e^{-\lambda z})^2} \sin(\psi - \psi_0), \tag{12}$$

where $\dot{\psi}_{\mathrm{T}}$ is the task-specific relative phase velocity (i.e., the system's desired task-level state velocity) that is a function of the current relative phase (i.e., task-space state).

Combining Eqs. (8), (9) and (12) for the window potential coupling forces, we derive the task-specific coupling terms at the component level:

$$\ddot{x}_i = (\omega_{0i} A_i \cos\phi_i) J_i^{+} \left(\frac{2\lambda e^{-\lambda z}}{(1 + e^{-\lambda z})^2} \sin[\psi - \psi_0] \right). \tag{13}$$

Finally, these task-specific coupling forces are added to the right-hand sides of the intrinsic component equations of motions (Eq. (2)), which are integrated numerically to generate position and velocity trajectories for each component oscillator.

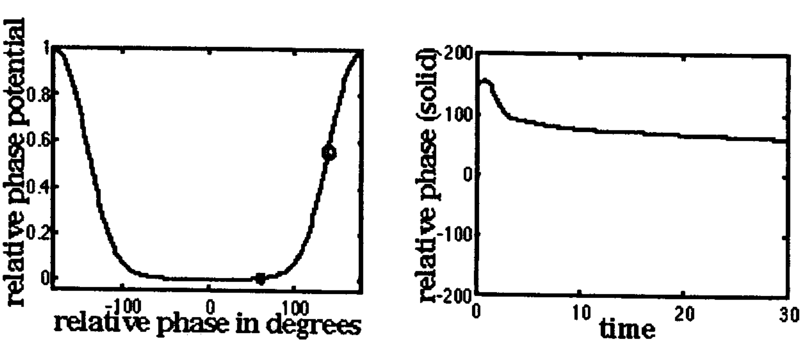

Fig. 4. Left: Task-space potential for window relative phasing. Right: Relative phase trajectory. 'o' = initial relative phase, '*' = final relative phase.

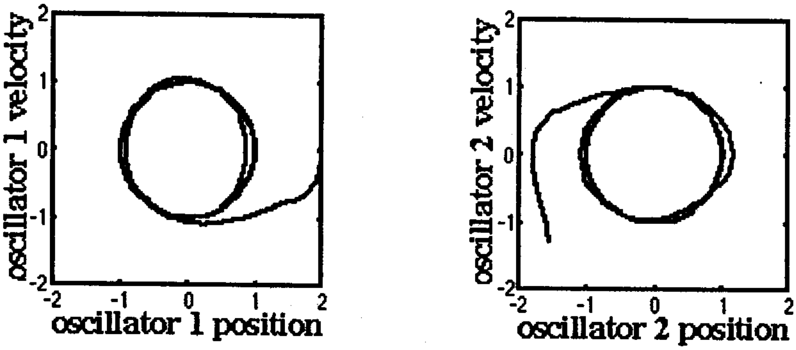

Fig. 5. Sinusoidal steady-state phase plane trajectories for oscillators 1 and 2, under point potential control.

3.3. *Relative phasing: Simulation results for punctate and window relative phasing*

The following simulations demonstrate the basic differences between point and window potentials, as well as the effects of startup transients and initial conditions. Parameter a for the point potential (Eq. (3)) and λ for the window potential (Eq. (11)) are set to 1 and 4, respectively; and window or point centers of the potential well (ψ_o) are set to 0°. All simulation durations are 30 s. Initial conditions are specified as follows: the initial amplitude and phase of oscillator 1 are fixed at 2 and 0°; the initial amplitude of oscillator 2 is fixed at 2 and its initial phase is set according to the desired initial relative phase. Intrinsic oscillator parameters (Eq. (2)) ω_{0i}, β_i, and γ_i are equal to 1, and $\alpha_i = -1$, resulting in sinusoidal steady-state motions of the uncoupled oscillators, that is, oscillations with a single frequency component evidenced by circular or elliptical orbits in the phase plane (e.g., Fig. 5) [3].

When initial relative phase is set to +140°, the resultant steady-state relative phase is *not* the same using the point and window potentials. With the point potential (Fig. 2, right panel), the final relative phase is nearly $0°(\psi_{\text{final}} = +0.0004°)$. However, with the window potential (Fig. 4, right

[3] Note that in contrast to these sinusoidal oscillations, if the ω_{0i}, α_i, β_i, and γ_i are varied independently, the uncoupled oscillators will generally display nonsinusoidal, *relaxation* oscillation trajectories that have more than one frequency component. When the oscillators are coupled, the system's relative phase does not settle smoothly toward its final value using either point or window potentials, but rather settles to an oscillatory steady state around its final value. (See also Fuchs, Jirsa, Haken and Kelso (1996) and Sternad, Turvey and Saltzman (1999c) for more extensive discussions of relative phase trajectories in coupled relaxation oscillators).

panel), the oscillators settle to a relative phase inside the window but not at its center, $\psi_{\text{final}} = +60°$. [4] In both these simulations relative phase increases transiently and then reverses direction and approaches its final value. The reason for this initial increase is that the total system behavior is due to the resultant combination of both the task-space coupling forces and intrinsic oscillator forces. While the task-space coupling acts to move the system monotonically from initial to target relative phase, the intrinsic forces are dependent on the system's initial conditions. In Figs. 2 and 4, the initial conditions cause the intrinsic forces to transiently dominate the system and push relative phase away from its final value.

Final relative phase, however, is critically affected by the nature of initial transients in the case of the window potential but is not particularly affected with a point potential. Thus, when initial relative phase is altered from +140° to −140° with the window potential, final relative phase moves from 60° to −54°, but it is unaffected with the point potential. Relatedly, when initial relative phase is already within the phase window, the transient behavior induced by the initial conditions may move the final relative phase farther from or nearer to the window center. Thus, $\psi_{\text{initial}} = 20°$ yields $\psi_{\text{final}} = 28°$, and $\psi_{\text{initial}} = -20°$ yields $\psi_{\text{final}} = -9°$.

3.4. An alternative means of capturing natural speech timing variability

The phase window approach has been put forward as a means of capturing the variability in relative timing that results from myriad influential factors in a natural speech setting. This approach has been argued to be superior to a punctate approach which does not *prima facie* allow for variability in the relative timing of two speech gestures. However, one can imagine a coupling relationship between two gestures that is punctate in nature but which crucially allows the minimum of the potential well to itself vary as a function of communicative factors [5]. This might be called the "wandering-well" approach and should be evaluated as an alternative to the phase window approach. While both the phase window and wandering-well approach require

[4] Although the point potential system has settled to its relative phase target value with a final velocity of zero at the end of the 30 s simulation, the relative phase of the window potential system is still slowly changing after 30 s. The reason for the window system's behavior is simply that the window potential's squashing function Eq. (11) does not create a perfectly flat potential floor, but rather leaves a very weakly curved surface centered at the target value.

[5] The possibility of such an alternative, punctate scenario was raised by an anonymous reviewer.

empirical data to define the range of variability permitted for particular gestural structures, the two approaches make different predictions regarding the distribution of tokens within that range of variability. Specifically, the phase window approach requires that tokens be distributed within a continuous range (although the distribution itself would not be expected to be uniform). Alternatively, the wandering-well approach allows for discontinuous distribution of observed timing relationships; for example, two gestures could be phased at precisely 0° under certain conditions and at precisely 240° under other conditions but an intervening phase relation of 200° might be impossible. Whether such multi-modal distributions of relative timing are found in natural speech articulatory data remains an open question. However, in our view continuous ranges of timing variability seem more plausible as these are the types of effects observed due to speaking rate and phrasal structure.

4. Modeling multifrequency rhythms

We have demonstrated in Section 3 how relative phase between two coupled oscillators can be flexibly controlled within a nonlinear dynamical systems framework such that target relative phase, the state variable at the task-space level, is defined as either a point or window. We view the phase window approach as advantageous for controlling the relative phasing of speech gestures, in that windows provide flexible control of intergestural timing and provide appropriately constrained variability as a function of linguistic and para-linguistic factors.

A further issue that must be addressed regarding the temporal patterning of speech is the creation of temporal order over increasingly extended and nested time spans, e.g., at the levels of syllables, words, and phrases. The issue of higher level order has been addressed in the speech domain by the *syllable-clock* hypothesis (e.g., Port, 1986; Port & Cummins, 1992) in which gestures are nested within syllable cycles. In turn, it has been hypothesized (O'Dell, 1995; O'Dell & Nieminen, 1999) that syllable-cycles are nested within cycles defined according to *feet* or *stress groups* (i.e., the intervals between stressed syllables or, more simply, interstress intervals [ISI's]) and, furthermore, that interstress intervals are nested within phrase-level *prosodic* cycles (Cummins & Port, 1998; Port, Cummins & Gasser, 1995; Tajima, Zawaydeh & Kitahara, 1999). Such nesting has been hypothesized to reflect the *n:1* frequency-locking dynamics of coupled nonlinear oscillators. For example, O'Dell & Nieminen (1999) explore the degree to which coupled syllable- and foot-level oscillators (n syllable cycles per 1 foot cycle) can account for cross-language variations in the temporally elastic relationship of foot duration to the average number of syllables per foot. Similarly, Cummins & Port (1998) hypothesize that the constrained set of phasing relationships found experimentally in their rhythmic speech tasks are the result of frequency locking between foot- and phrase-level oscillators.

In Section 3, we focused on the control of relative phasing between two 1:1 frequency-locked, coupled oscillators. In the present section, we show how the task-dynamic approach can be extended to controlling the frequency- and phase-locking characteristics of two coupled limit cycle oscillators for the production of multifrequency rhythms. In such rhythms, the frequency ratio ($\omega_1 : \omega_2$) between a pair of component oscillations can be reduced to m:n, where m and n are *relatively prime integers*, i.e., positive integers whose greatest common factor is 1 (e.g., Ogilvy and Anderson, 1966). Such rhythms repeat themselves with a global period, $T_G = mT_1 = nT_2$, where $T_i = 2\pi/\omega_i$ ($i = 1,2$) denotes the period of the i^{th} oscillator. When m or n equals 1, the rhythms are called *harmonic*. As just described, these types of rhythmic structures have been hypothesized to underlie the nested timescales of speech production. If neither m nor n equals 1, the rhythms are considered *dissonant*, *polyrhythmic*, or *cross-rhythmic*. Both harmonic and polyrhythmic structures have been the foci of a large number of studies, many of which are cited in discussions below, on the unimanual, bimanual, and multi-limb production of complex rhythms.

When either or both m and n do not equal 1, the measure of relative phase defined by Eq. (1) does not remain stationary over time for a given m:n pattern of oscillations, even in the case of perfect frequency locking. Rather, the definition of relative phase must be generalized according to

$$\psi_G = m\phi_2 - n\phi_1 (\text{mod } 2\pi), \tag{14}$$

where ψ_G is the system's *generalized relative phase* (e.g., Haken, Peper, Beek & Daffertshofer, 1996; Keith & Rand, 1984; Sternad, Turvey & Saltzman, 1999a); and m:n is the intended frequency ratio.

ψ_G is used as the task-space variable (order parameter) for modeling multifrequency rhythmic movements, and has the desired property of remaining constant over time when two oscillations with a frequency ratio $= \omega_1 : \omega_2 = m$:$n$ are offset relative to one another by a given temporal interval. For example, if each oscillator's waveform can be described as a monofrequency cosine function of time ($\cos[\omega_i t]$), then ψ_G remains constant over time at 0° whenever the relative temporal offset between waveforms results in the peaks of both waveforms ($\phi_i = 0$) being aligned at any point in time. Note that this condition implies that the waveforms' peaks become aligned periodically at the pattern's *global period*, T_G. Starting with the waveforms aligned such that $\psi_G = 0°$, as either waveform is continuously translated in time by $+(T_G/mn)$ (i.e., the waveform is phase delayed), ψ_G changes smoothly from 0° to −360°; as the waveform is continuously translated by $(-T_G/mn)$ (i.e., the waveform is phase advanced), ψ_G changes smoothly from 0° to +360°. Note that dividing the global period by mn creates a set of equal temporal subdivisions ($T_G/mn = T_1/n = T_2/m$) in a *rhythmic grid* (Weisberg, 1993) or *cross-rhythm* (Collier & Wright, 1995) that can be used to guide the learning and performance of $m{:}n$ polyrhythms (see also Fig. 1 in Deutsch, 1983).

Using Eq. (14), we defined a task-space control scheme for the punctate control of generalized relative phase by simply substituting ψ_G for ψ, and $\psi_{G,0}$ (desired target for ψ_G) for ψ_0, in Eqs. (3) and (4). By doing so we define a task-space potential and first-order task-space equation of motion:

$$V(\psi_G) = -a\cos(\psi_G - \psi_{G,0}), \tag{15}$$

$$\dot{\psi}_{G,T} = -dV/d\psi_G = -a\sin(\psi_G - \psi_{G,0}). \tag{16}$$

Similarly, we transformed the task-space velocity, $\dot{\psi}_{G,T}$, defined by Eq. (16) into a corresponding set of task-specific coupling terms for each oscillator's phase velocity by differentiating Eq. (14) with respect to time:

$$\dot{\psi}_G = m\dot{\phi}_2 - n\dot{\phi}_1 = J\dot{\tilde{\phi}}, \tag{17}$$

where $J = [-n, m]$, the 1×2 Jacobian matrix of Eq. (17) and $\dot{\tilde{\phi}} = [\dot{\phi}_2 - \dot{\phi}_1]^T$, the vector of phase velocities.

Inverting Eq. (17) yields

$$\dot{\tilde{\phi}} = J^*\dot{\psi}_G, \tag{18}$$

where $J^* = W^{-1}J^T(JW^{-1}J^T)^{-1}$, the 2×1 *weighted* pseudoinverse of the Jacobian, which provides a unique, weighted least squares solution for the set of computed component phase velocities used at each point in time; and

$$W = \begin{bmatrix} w_1 & 0 \\ 0 & w_2 \end{bmatrix}$$

is a positive definite matrix (typically diagonal) whose elements are constant during the production of a given single gesture, and each articulator weight, w_i, is associated with component i.

It is helpful to expand the formula for J^* into the following expression:

$$J^* = [-w_2 n, w_1 m]^{\mathrm{T}} \frac{1}{w_2 n^2 + w_1 m^2} = [J_1^*, J_2^*]. \tag{19}$$

Eqs. (17) and (18) can be combined to express the task-specific coupling terms for the ith oscillator's phase velocity:

$$\dot{\phi}_{i,\mathrm{T}} = J_i^* \dot{\Psi}_{\mathrm{G,T}}. \tag{20}$$

As shown earlier in the case of 1:1 punctate relative phasing, we use Eqs. (9) and (10) to transform each oscillator's task-specific phase velocity into a coupling term at the component level. Finally, we add these task-specific coupling forces to the right-hand sides of the components' equations of motion (Eq. (2)), and solve these equations numerically to generate state trajectories for the oscillators.

4.1. Simulation results: Examples

Using this method, one can generate stable polyrhythms with various frequency ratios and generalized relative phases. Fig. 6 illustrates the component oscillators' position trajectories for three $(m{:}n, \psi_{\mathrm{G},0})$ cases: $(2{:}1, 0°)$, $(3{:}2, 180°)$, and $(4{:}3, 90°)$. All simulations involve articulator weights $(w_i) = 1$, initial relative phases$= +140°$, simulation durations of 20 s, punctate task-space potential whose amplitude $(a) = 1$, and sinusoidal (mono-frequency) dynamics for the uncoupled component oscillators ($\alpha_i = -\omega_{0i}$, $\beta_i = \omega_{0i}, \gamma_i = 1/\omega_{0i}$). Although not included in Fig. 6, all generalized relative phase trajectories attained and remained at their respective target values after initial transients of approximately 4–6 s. As will be discussed further in Section 4.3, this figure demonstrates the ability of the model to generate performable $(m{:}n, \psi_{\mathrm{G},0})$ patterns (e.g., $[2{:}1, 0°]$), patterns that might be performable with sufficiently extended practice (e.g., $[3{:}2, 180°]$), and patterns that, if performable at all, would most likely require an inordinate amount of practice (e.g., $[4{:}3, 90°]$).

Additionally, one can examine the manner in which increasing the detuning of the component oscillator frequency ratio $(\omega_{0i}{:}\omega_{02})$ from the desired frequency ratio $(m{:}n = 2{:}1)$ changes the pattern from constant phase stability (Fig. 7(a)) at the target relative phase $(\psi_{\mathrm{G},0} = 0°)$, to periodic phase

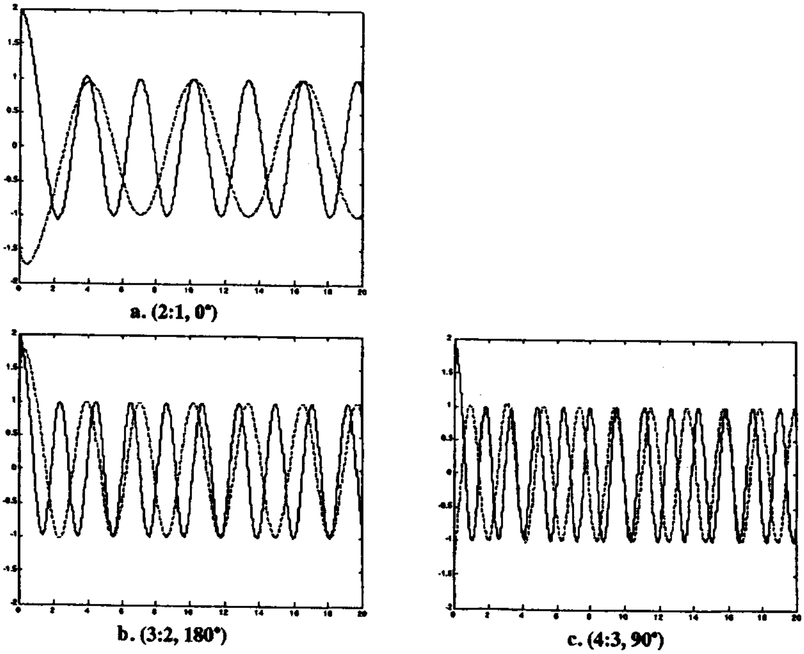

Fig. 6. Simulated polyrhythmic patterns with different $m{:}n$ frequency- and $\psi_{G,0}$ phase-lockings. (a) $(m{:}n, \psi_{G,0}) = (2{:}1, 0°), \omega_{01} = 2, \omega_{02} = 1$; (b) $(m{:}n, \psi_{G,0}) = (3{:}2, 180°), \omega_{01} = 3, \omega_{02} = 2$; (c) $(m{:}n, \psi_{G,0}) = (4{:}3, 90°), \omega_{01} = 4, \omega_{02} = 3$. Solid line = oscillator-1, broken line = oscillator-2. See text for further details.

stability (phase oscillation centered about 45°; Fig. 7(b)), to phase intermittency (center of phase oscillation changes slowly from 10° to 130° and then quickly jumps up to 10°; Fig. 7(c)), and to phase wandering or drift (Fig. 7(d)). Similar results of detuning-induced intermittency and phase wandering or drift in multifrequency bimanual rhythms have been reported in the literature (e.g. Kelso & DeGuzman, 1992).

4.2. Properties of the model: Role of the pseudoinverse

An analogy can be drawn between the present coupled oscillator model and the task dynamic model used for simulating unimanual reaching in nonspeech behavior (Saltzman & Kelso, 1987) or bilabial closing gestures in the case of speech (Saltzman & Munhall, 1989). Although reaching and speech gestures have been modeled using the discrete dynamics of point attractors, they share in common with the oscillator approach a set of two coordinate systems (component and task coordinates), a *forward mapping* from component to task level, and an *inverse mapping* from task to

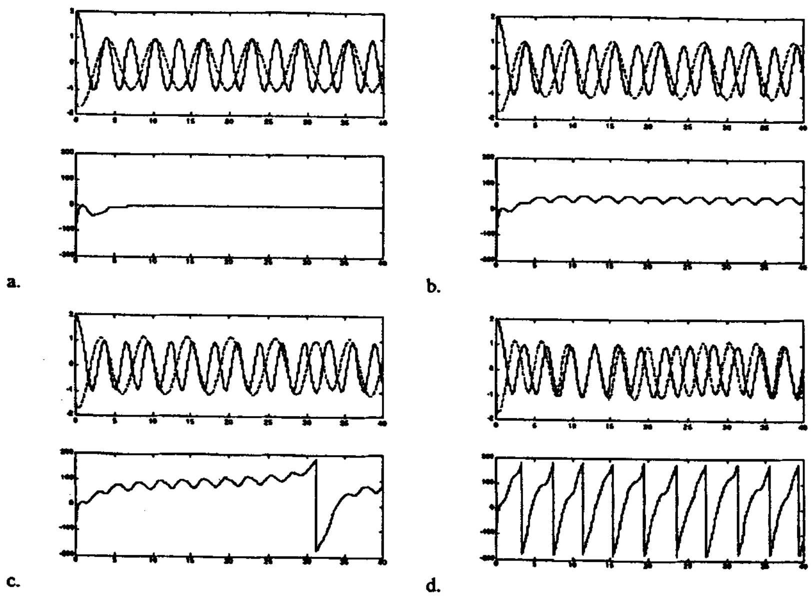

Fig. 7. Effects of increasing detuning of the components' frequency ratio ($\omega_{01}:\omega_{02}$) from the desired (m:n = 2:1) frequency ratio. Upper panel in each pair shows oscillator trajectories (solid line = oscillator-1, broken line = oscillator-2), lower panel shows generalized relative phase trajectory. (a) constant phase stability ($\omega_{01} = 2, \omega_{02} = 1$); (b) periodic phase stability ($\omega_{01} = 2, \omega_{02} = 1.25$); (c) phase intermittency, ($\omega_{01} = 2, \omega_{02} = 1.39$); (d) phase wandering/drift, ($\omega_{01} = 2, \omega_{02} = 1.9$). Simulation duration is 40s; all other parameter setting are as in Fig. 6(a).

component level (Fig. 3). Additionally, the oscillator work shares the property of *redundancy* with models of, for example, a reaching task in the horizontal plane where the x and y positions of the hand are controlled using three-arm joints (flexion/extension at the shoulder, elbow, and wrist), or of a bilabial closure task where the vertical distance between the lips is controlled using three articulators (jaw, upper and lower lips). In all cases there are fewer task degrees of freedom than articulator component degrees of freedom, and there are an infinite number of component positions that correspond to a given task-space position. This situation thus provides an adaptive flexibility in the manner in which trajectories unfold at the component level.

In the present modeling and in prior task-dynamic modeling of gestural patterning in speech production, the Jacobian pseudoinverse is used in the mapping from desired motion in the task space to a corresponding motion at the component level. In the coupled oscillator simulations, the pseudoinverse serves to partition the overall force defined in the task space into a set of phase-velocity coupling forces for each oscillator. For multifrequency oscillations ($m \neq n$, and assuming $m > n$, i.e., oscillator-1 is faster than oscillator-2 in the coupled state), the coupling force to the slow oscillator is greater than the coupling force to the fast oscillator. This result holds even in the case where the unweighted (or, equivalently, the identity-weighted) pseudoinverse is used, due to the different magnitudes of m and n (Eqs. (19) and (20); see also Haken et al. (1996) for related coupling term asymmetries). Thus, during the transients when the system is away from the target generalized relative phase, $\psi_{G,0}$, and hence the task level forces are nonzero, the task goal will be attained by the slow oscillator bearing the brunt of the task demands while the fast oscillator is relatively "insulated" from these task demands. This result of fast rhythms dominating slow rhythms is consistent with experimental findings on the performance of polyrhythms (e.g., Peper, Beek & Van Wieringen, 1995a,b; Peters & Schwartz, 1989; Sternad, Turvey & Saltzman, 1999b; Summers, Ford & Todd, 1993a, Summers, Rosenbaum, Burns & Ford, 1993b; cf. Klapp, Nelson & Jagacinski, 1998). An additional consequence of defining coupling in this manner for multifrequency rhythms is that the overall strength of coupling, proportional to $(|J_1^+| + |J_2^+|)$, decreases systematically with increasing values of m and n. If one assumes that the stability of m:n patterns is proportional to the overall strength of inter-oscillator coupling, then this result is consistent with the decreasing stability and increasing difficulty of higher order patterns compared to lower order ones (e.g., Collier & Wright, 1995; Deutsch, 1983; Jagacinski, Beek & Peper, 2000; Peper et al., 1995b; Summers et al., 1993b; Treffner & Turvey, 1993).

There are situations, however, in which the unweighted pseudoinverse is inadequate. For example, a given speaker may produce bilabial closing gestures using mostly lip motions while another speaker uses primarily jaw motion, due to habitual or strategic differences in speaking styles. Such interspeaker differences can be modeled using the weighted pseudoinverse, J^* (Eqs. (18) and (19)), according to which an articulator component receives a relatively larger (smaller) portion of the overall forcing defined in the task space if its corresponding weight, w_i, is relatively smaller (larger) than the other component's weight. Thus, if one component has a weight that is much larger than a second component's weight, the former component will dominate the latter. For modeling the movement patterns of coupled oscillators, the articulator weights of J^* may be useful in capturing coupling asymmetries other than those related to the task-specific frequency-locking behavior associated with parameters m and n. For example, the articulator weights may contribute to the inter-hand dominance effects associated with handedness and cognitive/attentional factors – preferred and/or attended hands typically display greater phase leads and less variability – that have been reported in studies of coupled rhythmic movements (e.g., Amazeen, Amazeen, Treffner & Turvey, 1997; Byblow, Bysouth-Young, Summers & Carson, 1998; Pressing, Summers & Magill, 1996; Treffner & Turvey, 1995).

4.3. Limitations of the model

In its current form, our model is limited in (at least) two respects. First, it is not powerful enough, in that it can display neither transitions from higher-order frequency lockings to lower-order lockings as performance rate is increased, nor spontaneous simplifications of higher-order timing patterns toward those associated with lower-order frequency lockings. Such transitions and simplifications have been reported experimentally (e.g., Byblow et al., 1998; Engbert, Scheffczyk, Krampe, Kurths & Kliegl, 1998; Peper et al., 1995b; Serrien & Swinnen, 1997; Summers & Pressing, 1994; Treffner & Turvey, 1993) and modeled theoretically (e.g., Engbert, Scheffczyk, Krampe, Rosenblum, Kurths & Kliegl 1997; Haken et al., 1996; Keith & Rand, 1984; Pressing, 1999). Our model does not display transitions primarily because it only contains a single coupling term that supports a particular *m:n* frequency-locking regime in contrast, for example, to Haken et al. (1996) who include a coupling term for every possible *m:n* case. Furthermore, the latter model includes a mechanism by which increases in performance rate serve to decrease the strengths of higher-order coupling terms more than the lower-order terms and, thereby, facilitate rate-conditioned transitions from higher to lower order polyrhythms. Our model does not include such frequency-dependent decreases in coupling strength. We hesitate to include all possible coupling terms because it is unclear whether they all need to be explicitly included. For example, a sinusoidally driven pendulum can be modeled by the *standard circle map* and displays the entire gamut of frequency-locking regimes with changes only in the driving frequency and strength of coupling between pendulum and driver (e.g., Baker & Gollub, 1996). Significantly, there is only one coupling term, yet all frequency locks are possible. It is thus possible that transitions among bimanual frequency-locks do not require explicit terms for each lock. Additionally, Haken et al. (1996) rely on amplitude decreases that accompany frequency increases to weaken the coupling coefficients as performance rate is quickened. This mechanism has since been challenged on empirical grounds (Peper & Beek, 1998a,b, 1999). Thus, we currently take a relatively conservative position regarding frequency-locking transitions and are exploring the possibility that such rate-induced transitions are due to rate-driven changes in: (a) the relationships between central driving frequencies and ongoing resonant frequencies of the biomechanical periphery, and in (b) the strengths of coupling between center and periphery.

Finally, the second limitation of our model is that it is too powerful and can generate frequency- and phase-locking patterns which are either impossible to produce or only produced after extremely long periods of training (see Fig. 6 and Section 4.1). Speech production and manual tapping tasks are in fact quite constrained in the patterns that are readily produced. For example, Semjen and Ivry (in press) demonstrated such constraints in a bimanual tapping task in which subjects attempted to produce a range of specified asynchronies between the hands. Using a continuation task paradigm, subjects tapped in time to an auditory template (synchronization task; different pitch tones for each hand) and then attempted to continue tapping the same temporal pattern after the tones had ceased (continuation task). Of interest was the discrepancy between the target phase and the phase produced by one hand inside the cycle of successive taps produced by the other hand. Subjects' tapping was clearly biased toward producing taps at phases of 1/2, 1/3, and 2/3. The authors interpret these results as providing additional support for the findings in the literature that it is easy to produce taps defining successive intervals with the ratios of 1:1 (phase = 1/2), 1:2 (phase = 1/3), or 2:1 (phase = 2/3), and that attempted productions of other higher order ratios are often distorted toward these simpler ratios (e.g., Fraisse, 1982). What is remarkable is that nearly identical findings have been reported for a speech production task (Cummins & Port, 1998; see also Port et al., 1995; Tajima et al., 1999) in which a phrase such as "*big* for a *duck*" (italics indicated the phrase's stressed syllables) was repeated rhythmically in time with an auditory metronome (high pitch tone). The metronome specified a succession of global cycles that began with each onset of the phrase-initial stressed syllable, and within each trial a low-pitch tone was presented at a given phase within the global cycle. Subjects were asked to synchronize the phrase-final syllable with this intervening tone, and a range of target phases were probed over the course of the experiment. Using the same type of synchronization-continuation paradigm as was used by Semjen and Ivry (in press), the same findings resulted for speech as for the tapping – namely, despite the goal of reproducing the task-demanded phasing patterns, speakers showed clear biases toward producing phasings of 1/2, 1/3, and 2/3. Such uniformity of results across manual and speech tasks suggests that a common dynamical structure underlies the production of these temporal patterns, and we are attempting to formulate such a structure as a system of coupled nonlinear oscillators. Thus, our search for behavioral timing constraints that emerge from the dynamics of an oscillatory network echoes the work of Golubitsky, Stewart, Buono and Collins (1998, 1999), who require that their locomotor central pattern generator (CPG) models generate the full set of an animal's naturally produced gaits and no others.

5. Concluding remarks

The coupled oscillator task dynamic method outline above (and represented schematically in Fig. 3) can be summarized as follows: (a) define a forward mapping of component state variables to task-space state variables, (b) define the desired task-specific dynamics (task-space state-velocity) as a function of current task-space state, (c) create an inverse map of task-space velocity to component accelerations, and (d) solve the resulting equations of motion at the component level. The consequent motion patterns will be the result of both the intrinsic dynamics of the oscillatory components and the task-specific coupling terms. This approach was explored in the context of two different considerations of interoscillator relative phase as a control variable – the implementation of a *phase window* in which interoscillator timing is specified as a range of relative phase; and a model of the control of the frequency- and phase-locking characteristics of coupled limit cycle oscillators for the production of multifrequency rhythms.

Using this method, intuitions can be gained as to the manner in which intrinsic oscillator waveform shapes (e.g., the sinsusoidal vs relaxation waveforms of the uncoupled oscillators noted in Footnote 2), detuning (differences between component intrinsic frequencies), task-specific frequency ratios (values of m and n), and other aspects of coupling structure (articulator weights) interact in the creation of coupled rhythmic patterns (resultant oscillation frequencies, interoscillator phasing, and waveform shapes). In particular, the resultant transient and steady-state trajectories of both phase and amplitude (or, equivalently, position and velocity) can be examined and used to generate detailed predictions that would not be possible using models derived with averaging techniques that, by definition, ignore the within-cycle temporal variation of these variables' trajectories (e.g., Haken et al., 1996; Kopell, 1988).

We conclude that the task-dynamic approach offers a relatively straightforward way to define coupling terms for any hypothesized set of component oscillators. This allows the user to explore conditions under which the coupled system can display desired stable patterns of interoscillator frequency locking and phase entrainment. Furthermore, this framework is not limited to simulating systems of coupled oscillators. It applies more generally to modeling processes of control and coordination in skilled activities of both the limbs and speech articulators.

Acknowledgements

The authors thank Michael O'Dell, Bruno Repp, Dagmar Sternad, and two anonymous reviewers for their insightful comments on several of the issues raised in this paper. The authors gratefully acknowledge the support of NIH grants DC-03663 and DC-03172, and the Santa Fe Institute's funding of the "Debates in Dynamics" workshop at Pennsylvania State University (August, 1999) organized by Fellow-at-Large Dagmar Sternad.

References

Abry, C. & Lallouache, T. (1995). Modeling lip constriction anticipatory behavior for rounding in French with MEM (Movement Expansion Model). In K. Elenius, & P. Branderud, *Proceedings of the XIIth International Congress of Phonetic Sciences* (pp. 152–155). Stockholm: KTH and Stockholm University.

Amazeen, E. L., Amazeen, P. G., Treffner, P. J., & Turvey, M. T. (1997). Attention and handedness in bimanual coordination dynamics. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 1552–1560.

Baker, G. L., & Gollub, J. R. (1996). *Chaotic dynamics: An introduction* (2nd ed.). New York: Cambridge University Press.

Bell-Berti, F., & Harris, K. (1981). A temporal model of speech production. *Phonetica, 38*, 9–20.

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

Bullock, D., & Grossberg, S. (1988). Neural dynamics of planned arm movements: Emergent invariants and speed-accuracy properties during trajectory formation. *Psychological Review, 95*, 49–90.

Byblow, W. D., Bysouth-Young, D., Summers, J. J., & Carson, R. G. (1998). Performance asymmetries and coupling dynamics in the acquisition of multifrequency bimanual coordination. *Psychological Research, 61*, 56–70.

Byrd, D. (1996a). A phase window framework for articulatory timing. *Phonology, 13*, 139–169.

Byrd, D. (1996b). Influences on articulatory timing in consonant sequences. *Journal of Phonetics, 24*, 209–244.

Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (2000). Phrasal signatures in articulation. In M. B. Broe, & J. B. Pierrehumbert (Eds.). *Papers in Laboratory Phonology V* (pp. 70–87). Cambridge: Cambridge University Press.

Byrd, D., & Saltzman, E. (1998). Intragestural dynamics of multiple phrasal boundaries. *Journal of Phonetics, 26*, 173–199.

Byrd, D., & Tan, C. C. (1996). Saying consonant clusters quickly. *Journal of Phonetics, 24,* 263–282.

Coker, C. H. (1976). A model of articulatory dynamics and control. *Proceedings of the IEEE, 64,* 452–460.

Collier, G. L., & Wright, C. E. (1995). Temporal rescaling of simple and complex ratios in rhythmic tapping. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 602–607.

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics, 26,* 145–171.

Deutsch, D. (1983). The generation of two isochronous sequences in parallel. *Perception and Psychophysics, 34,* 331–337.

Engbert, R., Scheffczyk, C., Krampe, R. T., Kurths, J., & Kliegl, R. (1998). Symbolic dynamics of bimanual production of polyrhythms. In H. Kantz, J. Kurths, & G. Mayer-Kress (Eds.). *Nonlinear analysis of physiological data* (pp. 271–282). New York: Springer.

Engbert, R., Scheffczyk, C., Krampe, R. T., Rosenblum, M., Kurths, J., & Kliegl, R. (1997). Tempo-induced transitions in polyrhythmic hand movements. *Physical Review E, 56,* 5823–5833.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). New York: Academic Press.

Fuchs, A., Jirsa, V. K., Haken, H., & Kelso, J. A. S. (1996). Extending the HKB model of coordinated movement to oscillators with different eigenfrequencies. *Biological Cybernetics, 74,* 21–30.

Golubitsky, M., Stewart, I., Buono, P.-L., & Collins, J. J. (1998). A modular network for legged locomotion. *Physica D, 115,* 56–72.

Golubitsky, M., Stewart, I., Buono, P.-L., & Collins, J. J. (1999). Symmetry in locomotor central pattern generators and animal gaits. *Nature, 401,* 693–695.

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

Haken, H., Peper, C. E., Beek, P. J., & Daffertshofer, A. (1996). A model for phase transitions in human hand movements during multifrequency tapping. *Physica D 90,* 179–196; *92,* 260 erratum.

Hardcastle, W. J. (1985). Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication, 4,* 247–263.

Henke, W. (1966). Dynamic articulatory models of speech production using computer simulation. Unpublished Ph.D. dissertation, MIT, Cambridge, MA.

Jagacinski, R. J., Beek, P. J., & Peper, C. E. (2000). Dynamic, stochastic, and topological aspects of polyrhythmic performance. *Journal of Motor Behavior, 32,* 323–326.

Kay, B. A., Kelso, J. A. S., Saltzman, E. L., & Schöner, G. (1987). Space-time behavior of single and bimanual rhythmical movements: Data and limit cycle model. *Journal of Experimetal Psychology: Human Perception and Performance, 13,* 178–192.

Keith, W. L., & Rand, R. H. (1984). 1:1 and 2:1 phase entrainment in a system of two coupled limit cycle oscillators. *Journal of Mathematical Biology, 20,* 133–152.

Kelso, J. A. S., & DeGuzman, G. C. (1992). The intermittent dynamics of coordination. In G. E. Stelmach, & J. Requin (Eds.), *Tutorials in motor behavior II* (pp. 549–561). New York: Elsevier.

Kelso, J. A. S., Saltzman, E. L., & Tuller, B. (1986a). The dynamical perspective on speech production: Data and theory. *Journal of Phonetics, 14,* 29–60.

Kelso, J. A. S., Saltzman, E. L., & Tuller, B. (1986b). Intentional contents, communicative context, and task dynamics: A reply to the commentators. *Journal of Phonetics, 14,* 171–196.

Klapp, S. T., Nelson, J. M., & Jagacinski, R. J. (1998). Can people tap concurrent bimanual rhythms independently? *Journal of Motor Behavior, 30,* 301–322.

Kopell, N. (1988). Toward a theory of modelling central pattern generators. In A. H. Cohen, S. Rossignol, & S. Grillner (Eds.), *Neural control of rhythmic movements in vertebrates* (pp. 369–413). New York: Wiley.

Kröger, B. J., Schröder, G., & Opgen-Rhein, C. (1995). A gesture-based dynamic model describing articulatory movement data. *Journal of the Acoustical Society of America, 98,* 1878–1889.

Krakow, R. A. (1993). Nonsegmental influences on velum movement patterns: Syllables, sentences, stress, and speaking rate. In M. K. Huffman, & R. A. Krakow (Eds.), *Phonetics and phonology, vol. 5. Nasals, nasalization, and the velum* (pp. 87–116). New York: Academic Press.
Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science, 6*, 177–208.
Lubker, J. (1981). Temporal aspects of speech production: Anticipatory labial coarticulation. *Phonetica, 38*, 51–65.
Lubker, J. (1986). Articulatory timing and the concept of phase. *Journal of Phonetics, 14*, 133–137.
Munhall, K. G., Löfqvist, A., & Kelso, J. A. S. (1994). Lip-larynx coordination in speech: Effects of mechanical perturbations to the lower lip. *Journal of the Acoustical Society of America, 95*, 3605–3616.
Nittrouer, S. (1991). Phase relations of jaw and tongue tip movements in the production of VCV utterances. *Journal of the Acoustical Society of America, 90*, 1806–1815.
Nittrouer, S., Munhall, K. G., Kelso, J. A. S., Tuller, B., & Harris, K. S. (1988). Patterns of interarticulator phasing and their relation to linguistic structure. *Journal of the Acoustical Society of America, 84*, 1653–1661.
Nguyen, M. H., & Cottrell, G. W. (1997). Tau-net: A neural network for modeling temporal variability. *Neurocomputing, 15*, 249–271.
O'Dell, M. L. (1995). *Intrinsic timing in a quantity language*. Licentiate Dissertation, Department of Linguistics, University of Jyväskylä, Finland.
O'Dell, M. L., & Nieminen, T. (1999). Coupled oscillator model of speech rhythm. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences, vol. 2* (pp. 1075–1078). New York: American Institute of Physics.
Ogilvy, C. S., & Anderson, J. T. (1966). *Excursions in number theory*. New York: Dover.
Ostry, D. J., Gribble, P., & Gracco, V. L. (1996). Coarticulation of jaw movements in speech production: Is context sensitivity in speech kinematics centrally planned? *Journal of Neuroscience, 16*, 1570–1579.
Peper, C. E., & Beek, P. J. (1998a). Are frequency-induced transitions in rhythmic coordination mediated by a drop in amplitude? *Biological Cybernetics, 79*, 291–300.
Peper, C. E., & Beek, P. J. (1998b). Distinguishing between the effects of frequency and amplitude on interlimb coupling in tapping a 2:3 polyrhythm. *Experimental Brain Research, 118*, 78–92.
Peper, C. E., & Beek, P. J. (1999). Modeling rhythmic interlimb coordination: The roles of movement amplitude and time delays. *Human Movement Science, 18*, 263–280.
Peper, C. E., Beek, P. J., & Van Wieringen, P. C. W. (1995a). Coupling strength in tapping a 2:3 polyrhythm. *Human Movement Science, 14*, 217–245.
Peper, C. E., Beek, P. J., & Van Wieringen, P. C. W. (1995b). Multifrequency coordination in bimanual tapping: Asymmetrical coupling and signs of supercriticality. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1117–1138.
Perkell, J. S., & Matthies, M. L. (1992). Temporal measures of labial coarticulation for the vowel /u/: Within- and cross-subject variability. *Journal of the Acoustical Society of America, 91*, 2911–2925.
Peters, M., & Schwartz, S. (1989). Coordination of the two hands and effects of attention manipulation in the production of a bimanual 2:3 polyrhythm. *Australian Journal of Psychology, 41*, 215–224.
Port, R. (1986). *Translating linguistic symbols into time* (pp. 156–173). (Research in Phonetics and Computational Linguistics, Report No. 5), Department of Linguistics and Computational Linguistics, Indiana University, Bloomington, IN.
Port, R. F., & Cummins, F. (1992). The English voicing contrast as velocity perturbation. In *Proceedings of the 1992 International Conference on Spoken Language Processing (ICSLP '92), vol. 2 (pp. 1311–1314), Banff, Alberta, Canada*. Edmonton, Canada: Priority Printing.
Port, R., Cummins, F., & Gasser, M. (1995). *A dynamic approach to rhythm in language: Toward a temporal phonology*. (Technical Report No. 105), Indiana University Cognitive Science Program, Bloomington, IN.
Pressing, J. (1999). The referential dynamics of cognition and action. *Psychological Review, 106*, 714–747.

Pressing, J., Summers, J., & Magill, J. (1996). Cognitive multiplicity in polyrhythmic pattern performance. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1127–1148.

Rubin, P., Saltzman, E., Goldstein, L., McGowan, R., Tiede, M., & Browman, C. (1996). CASY and extensions to the task-dynamic model. In *Proceedings of the 1st ESCA Tutorial.*

Saltzman, E. (1995). Intergestural timing in speech production: Data and modeling. In K. Elenius, & P. Branderud (Eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences* (pp. 84–91). Stockholm: KTH and Stockholm University.

Saltzman, E., & Byrd, D. (1999). Dynamical simulations of a phase window model of relative timing. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey, (Eds.). *Proceedings of the XIVth International Congress of Phonetic Sciences* (pp. 2275–2278). New York: American Institute of Physics.

Saltzman, E. L., & Kelso, J. A. S. (1987). Skilled actions: A task dynamic approach. *Psychological Review, 94*, 84–106.

Saltzman, E., Löfqvist, A., Kay, B., Kinsella-Shaw, J., & Rubin, P. (1998). Dynamics of intergestural timing: A perturbation study of lip-larynx coordination. *Experimental Brain Research, 123*, 412–424.

Saltzman, E., Löfqvist, A., & Mitra, S. (2000). Clocks and glue: Global timing and intergestural cohesion. In M. B. Broe, & J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V* (pp. 88–101). Cambridge: Cambridge University Press.

Saltzman, E., & Mitra, S. (1998). A task-dynamic approach to gestural patterning in speech: A hybrid recurrent network. *Journal of the Acoustical Society of America 103 (5; Pt.2)*, 2893 (Abstract).

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

Semjen, A., & Ivry, R. B. (in press). The coupled oscillator model of between-hand coordination in alternate-hand tapping: A reappraisal. Journal of Experimental Psychology: Human Perception and Performance.

Serrien, D. J., & Swinnen, S. P. (1997). Coordination constraints induced by effector combination under isofrequency and multifrequency combinations. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 1493–1510.

Shaiman, S., Adams, S. G., & Kimelman, M. D. Z. (1995). Timing relationships of the upper lip and jaw across changes in speaking rate. *Journal of Phonetics, 23*, 119–128.

Shaiman, S., & Porter, R.J., Jr. (1991). Different phase-stable relationships of the upper lip and jaw for production of vowels and dipthongs. *Journal of the Acoustical Society of America, 90*, 3000–3007.

Sproat, R., & Fujimura, O. (1993). Allophonic variation in English/l/ and its implications for phonetic implementation. *Journal of Phonetics, 21*, 291–312.

Sternad, D., Turvey, M. T., & Saltzman, E. L. (1999a). Dynamics of 1:2 coordination: Generalizing relative phase to *n:m* rhythms. *Journal of Motor Behavior, 31*, 207–223.

Sternad, D., Turvey, M. T., & Saltzman, E. L. (1999b). Dynamics of 1:2 coordination: Sources of symmetry breaking. *Journal of Motor Behavior, 31*, 224–235.

Sternad, D., Turvey, M. T., & Saltzman, E. L. (1999c). Dynamics of 1:1 coordination: Temporal scaling, latent 1:1 and bistability. *Journal of Motor Behavior, 31*, 236–247.

Summers, J. J., Ford, S., & Todd, J. A. (1993a). Practice effects on the coordination of the hands in a bimanual tapping task. *Human Movement Science, 12*, 111–133.

Summers, J. J., Rosenbaum, D. A., Burns, B. D., & Ford, S. K. (1993b). Production of polyrhythms. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 416–428.

Summers, J. J., & Pressing, J. (1994). Coordinating the two hands in polyrhythmic tapping. In S. P. Swinnen, H. Heuer, J. Massion, & P. Casaer (Eds.), *Interlimb coordination: Neural, dynamical, and cognitive constraints* (pp. 571–593). New York: Academic Press.

Tajima, K., Zawaydeh, B. A., & Kitahara, M. (1999). A comparative study of speech rhythm in Arabic, English, and Japanese. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences, vol. 1* (pp. 285–288). New York: American Institute of Physics.

Treffner, P. J., & Turvey, M. T. (1993). Resonance constraints on polyrhythmic movements. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 1221–1237.

Treffner, P. J., & Turvey, M. T. (1995). Handedness and the asymmetric dynamics of bimanual rhythmic coordination. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 318–333.

Tuller, B., & Kelso, J. A. S. (1991). The production and perception of syllable structure. *Journal of Speech and Hearing Research, 34*, 501–508.

Weisberg, A. (1993). *Performing twentieth-century music: A handbook for conductors and instrumentalists*. New Haven, CT: Yale University Press.