Tutorial on task dynamics and articulatory phonology

Elliot Saltzman

Boston University, Boston, MA;

Haskins Laboratories, New Haven CT;

Contact Information:
(esaltz@bu.edu)

***Acknowledgements**: This work was supported by NIH grant R21-DC-009466 and NSF grant IIS-0703782;*

TISLR_Purdue_Oct'10

Colleagues

Martha Tyrone

Long Island University, Brooklyn NY

Haskins Laboratories, New Haven CT

Louis Goldstein

University of Southern California, Los Angeles CA

Haskins Laboratories, New Haven CT

Hosung Nam

Haskins Laboratories, New Haven CT

Gaurav Mathur

Gallaudet University, Washington DC

TISLR_Purdue_Oct'10

Overview of Presentation

- General Introduction:
  - Articulatory phonology: linguistics
  - Task-dynamics: movement science
- Review: Task-dynamic model of sensorimotor coordination
- Review: Articulatory phonology and the task-dynamic model of speech production
  - Gestural (spoken) and prosodic aspects
- Progress & challenges: Toward an articulatory phonology and task-dynamic model of signing

TISLR_Purdue_Oct'10

Part I. Articulatory Phonology: Major Hypotheses

Catherine Browman and Louis Goldstein

- Speech can be described in a way that captures its phonological and physical properties in a unitary structure .
- Act of speaking can be decomposed into atomic units, or vocal tract ***gestures***.
  - Units of information: Linguistically contrastive primitives of speech production
  - Units of action: Control structures that govern the creation and release of constrictions by distinct vocal tract organs (e.g., lips, tongue tip, tongue body)
  - **Coordinated** into larger 'molecular' lexical structures

TISLR_Purdue_Oct'10

Task Dynamics: Major Hypotheses

- Domain independence
  - The principles underlying skilled actions of the limbs, head, and torso are the *same* as those involved in the control and coordination of the speech articulators
- Unitary treatment of underlying invariance of representation and surface variability of performance
  - Motor equivalence: multiple articulator trajectories accomplish single goal
  - Underlying invariance: dynamics (force fields)
  - Surface variability: contextually varying kinematics (motion patterns) that emerge from the dynamics

TISLR_Purdue_Oct'10

Part II. Task Dynamics—4 Issues

Main Question: How can we build a "simple" (few degrees-of-freedom) special purpose device out of a many degree-of-freedom musculoskeletal system?

- 1st: Define the appropriate End-Effector for the task
  - Definition: A task's end-effector is defined in relation to the body part/parts whose motion most directly defines the task's special purpose device
  - Ex) hand in a reaching/punching task
  - Ex) forehead for heading a soccer ball
  - Ex) both hands for clapping/applause

TISLR_Purdue_Oct'10

Task Dynamics: 4 Issues (cont.)

- 2nd: Define the two *sets* of variables (coordinate spaces) that are needed to control the intended action
  - *dimensionalities* of these spaces
- Two important coordinate spaces are task space and articulator/musculoskeletal space
  - Task space: low dimensional; where the end-effector "lives"
    - Example: Reaching—3 translational degrees of freedom
    - Example: Wine glass transport—3 translational and 3 rotational degrees of freedom
  - Articulator/musculoskeletal space: higher dimensional; where the actuators "live"
    - Examples: joint angles, segment orientation angles

TISLR_Purdue_Oct'10

Task Dynamics: 4 Issues (cont.)

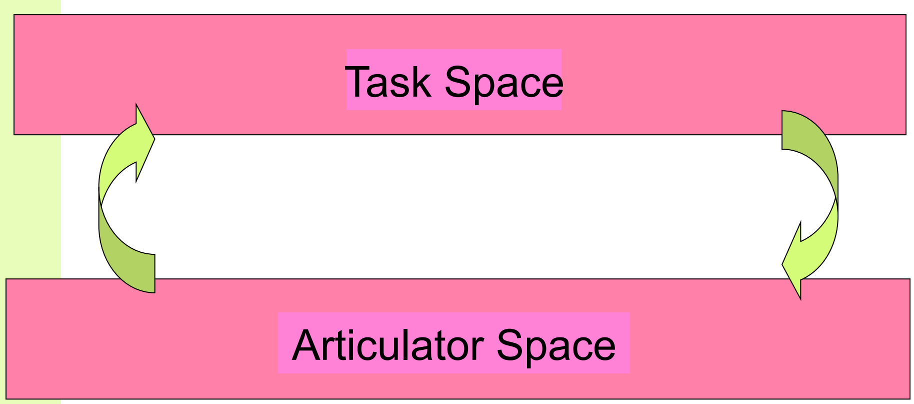
- 3rd: Define the appropriate *kind* of task-specific **dynamics** for the end-effector's degrees-of-freedom in the task space
  - Ex) Point attractor dynamics for discrete targeting/ positioning tasks
    - Reaching/punching: hand moves to contact target
    - Heading soccer ball: forehead moves to hit incoming ball
  - Ex) Limit cycle dynamics for sustained rhythmic tasks
    - Clapping/applause: both hands move rhythmically and symmetrically relative to one another
    - Polishing/scratching: hand/finger moves rhythmically on a surface

TISLR_Purdue_Oct'10

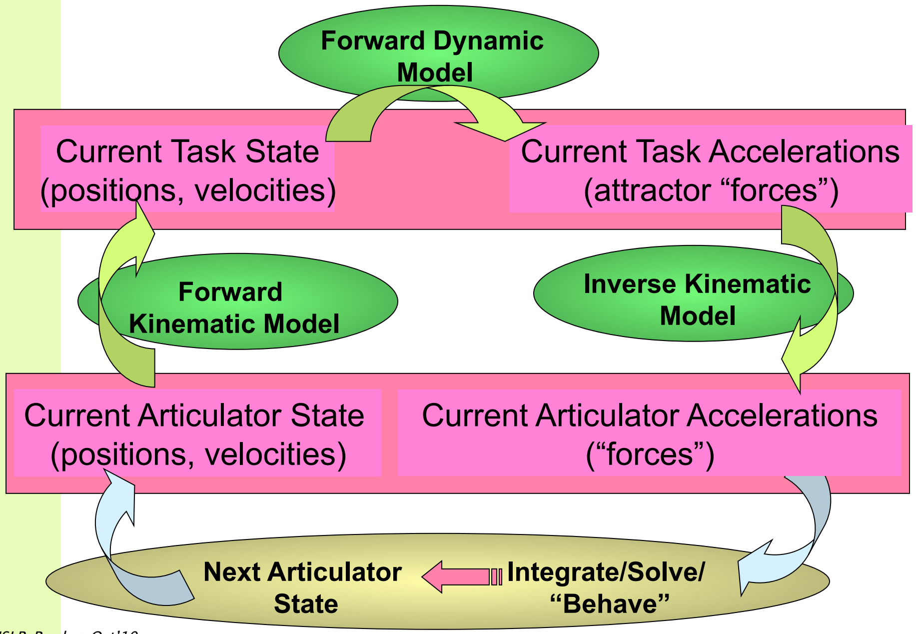
Task Dynamics: 4 Issues (cont.)

- 4th: Convert the task-specific accelerations on the end-effector into a pattern of corresponding joint accelerations in the articulator degrees-of-freedom
  - Harness the articulators in order to produce task-specific patterns of coordinated motion
    - need to know the kinematic/geometric relationships between the articulator and end-effector coordinate systems
  - These relationships are bidirectional, and include both:
    - Forward Kinematics: perceptual input transformed into task-relevant form
      - Ex)) joint angles are transformed into position and orientation of hand
    - Inverse Kinematics: Desired end-effector motion is transformed into required articulator motion
      - Ex) Desired hand acceleration converted into required task-specific, pattern of joint angular accelerations

Schematic of Task-Dynamic Model



Schematic of Task-Dynamic Model



Part III
Speech Production: An Articulatory Phonology and Task-Dynamic Model

Speech Gestures

- Definition: Equivalence classes of goal-directed actions by different sets of articulators in the vocal tract
  - examples:
    - bilabial gestures /p/, /b/, /m/—Upper lip, lower lip, and jaw work together to close and open the lips.
    - vocalic gestures /a/, /o/—Tongue body and jaw work together to position and shape the tongue dorsum (surface) for the vowel.
- Articulatory Phonology (Browman & Goldstein)
  - Gestures are the basic functional units of speech ("atoms"); syllables/words are "molecules"

TISLR_Purdue_Oct'10

Components of the Task-Dynamic Model

- There are 2 main components in the task dynamic model of speech production
  - Constriction formation/release component
    - shapes articulator trajectories given gestural timing information as input
    - Uses *tract-variable (constriction)* and *model articulator* coordinates.
  - Planning component: provides a dynamics of gestural timing
    - Uses *activation* and *planning oscillator ("clock")* coordinates

TISLR_Purdue_Oct'10

Task-Dynamics: Constriction (Tract-variable) Task Space & Model Articulator Space Coordinates

	Tract variables	Model articulators
LP	lip protrusion	upper & lower lips
LA	lip aperture	upper & lower lips, jaw
TDCL	tongue dorsum constrict location	tongue body, jaw
TDCD	tongue dorsum constrict degree	tongue body, jaw
LTH	lower tooth height	jaw
TTCL	tongue tip constrict location	tongue tip, body, jaw
TTCD	tongue tip constrict degree	tongue tip, body, jaw
TTCO	tongue tip constrict orientation	tongue tip, body, jaw
VEL	velic aperture	velum
GLO	glottal aperture	glottal width
$\mathbf{P_S}$	subglottal pressure	total lung force
$\mathbf{P_T}$	transglottal pressure	supralaryngeal vocal tract volume
$\mathbf{\Delta F_{OV}}$	delta virtual fundamental frequency	vocal fold tension, total lung force glottal width

Task-Space Dynamics of Individual Speech Gestures

- Task-space dynamics of single speech gestures are modeled using the *point attractor* dynamics of damped mass-spring systems
  - Examples: closing or opening the lips, raising or lowering the tongue tip, etc.
- Self-organized motor equivalence
  - Task-space: different initial conditions result in different paths to attain a given constriction target
  - Articulator-space: if an articulator is perturbed along the way, the others instantaneously and automatically compensate

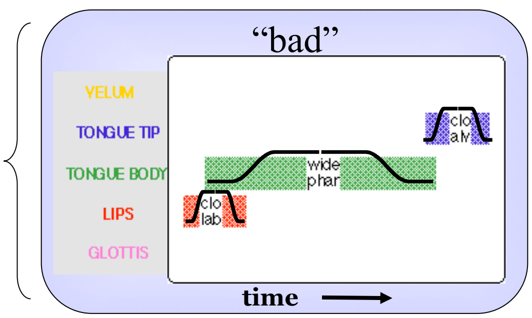
Planning Component: Gestural Activation

- A given gesture's constriction-dynamics influence vocal tract activity for a discrete interval of time
  - Activation interval
    - Activations wax and wane gradually at edges.
  - A gesture's strength is defined by its activation level (range: 0-1)
- In a given utterance, inter-gestural timing is determined by how the activation waveshapes of the component gestures evolve over time
  - Activation timing is controlled by a "clock" defined by a set of planning oscillators

TISLR_Purdue_Oct'10

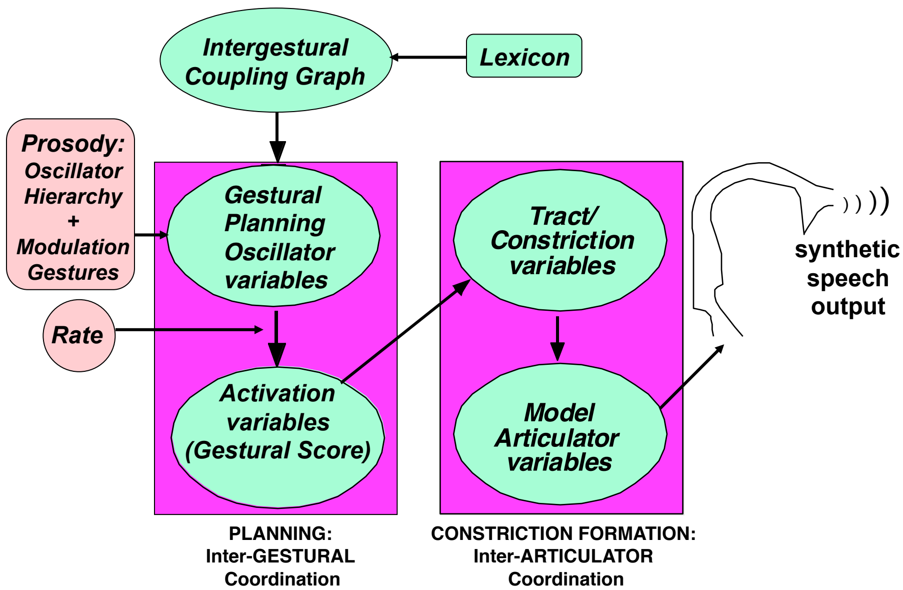
Gestural Scores

- The activation trajectories for the set of gestures in a given utterance is described by a gestural score
  - Rows = task-space (constriction) variables
  - Each row: activation vs. time waveforms for the associated task-space (constriction) variable
- Gestural Score



TISLR_Purdue_Oct'10

Task-Dynamic Model: Summary



TISLR_Purdue_Oct'10

Prosodic Effects of Phrasal Boundaries:

- Phrasal boundary effects on articulation include:
  - lengthening of gestural durations
  - decreased overlap (coarticulation) between adjacent gestures
  - spatially larger gestures in phrase-initial positions
- Boundary effects appear to be graded
  - stronger boundaries induce greater lengthening
- Question: How can we account for the variations of gestural timing associated with prosodic context?

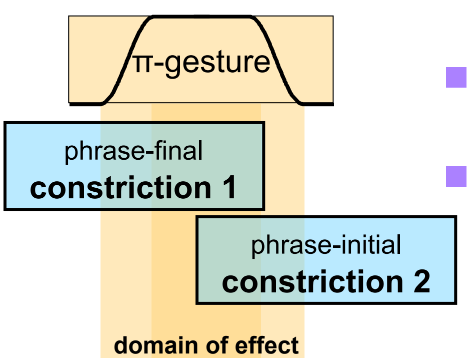
TISLR_Purdue_Oct'10

How is this Prosodic Action Effected?— Central Clock Slowing

- Hypothesis: Prosodic effects are induced by time slowing at the gestural control level.
  - slowing the timecourse of gestural activation
  - influence the expression of all constriction gestures spanning a phrasal boundary
- Slowing the central clock has the desired effects on both within-gesture and between-gesture timing.
  - gestural lengthening, spatial strengthening, reduced intergestural overlap

TISLR_Purdue_Oct'10

Juncture as Prosodic (π)- Gesture



Byrd, Kaun, Naryanan, & Saltzman (2000), Byrd (2000), Byrd & Saltzman (2003)

- Two constriction gestures overlap at a phrase boundary
- π-gesture straddles the phrase boundary.
- Constriction gestures are slowed during the activation interval of the π-gesture
  - degree of slowing is proportional to π-gesture's activation level
  - activation level determined by boundary strength

TISLR_Purdue_Oct'10

Part IV
Toward an Articulatory Phonology and Task-Dynamic Model of Signing

Sign Modeling: Extrapolations Consistent with Current Speech Model

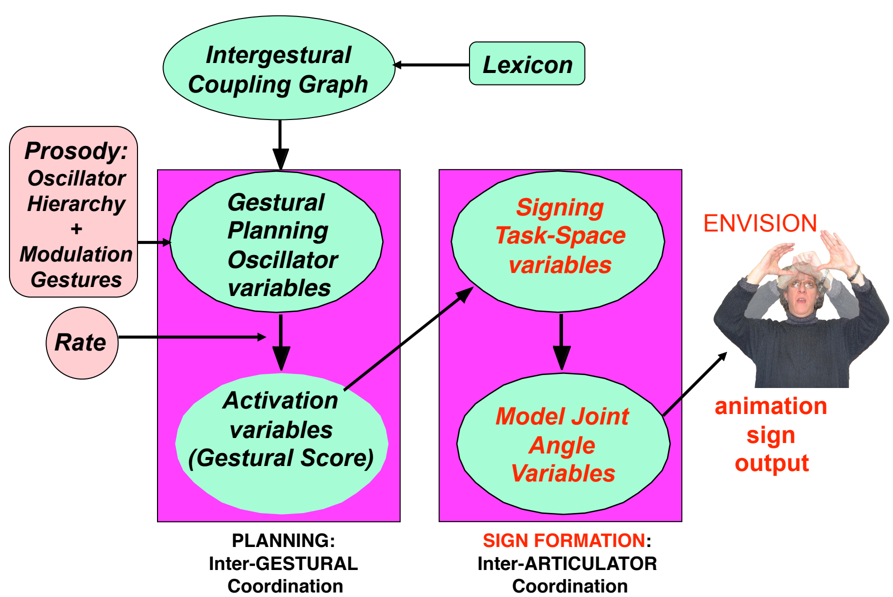
- Signs as multidimensional lexical objects
  - Sublexical dimensions defined by task spaces
    - Location, orientation, handshape, pathshape (movement)
- Hierarchy of functional units
  - Gestures: Task-space "atomic" units of action (contrastive phonological primitives)
  - Signs: "molecules" of gestures
- Gestural scores
  - Gestural activation trajectories for each task-space in a given utterance
- Phrasal boundary-induced lengthening
  - π-gestures

Sign Modeling: Properties Not Present in Current Model

- Articulatory phonology:
  - New types of primitive 'objects' (limit cycles)
    - Ex) rhythmic hand/finger motion in ASL: TREE
  - Multiphase units more complex than constriction formation (gap closing) → release (gap opening)
    - Ex) motion to target region → handshape change →motion away from target (ASL: STUBBORN)
- Task dynamics
  - Motion in 3-dimensional space, not 2-D midsagittal plane
  - Target complexity
    - Areas on body surface(s), not points; Volumes in space, not points
    - Coordinate system axes (relative attractive strengths) "embedded" in target areas/volumes
  - Task-space distances between surfaces and points or other surfaces (not point-to-point)
  - Body as complex, moving spatial array of obstacles (repellers) and time-varying targets (attractors)

TISLR_Purdue_Oct'10

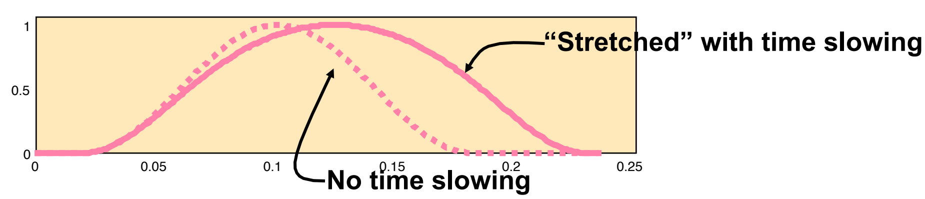
Planned Task-Dynamic Model of Signing



TISLR_Purdue_Oct'10

Thank you

Slowing activation timecourse



Equation for time scaling/stretching/slowing:

$$\tau(t) \xrightarrow{d/dt} \dot{\tau} = \frac{d}{dt}\tau = (1 - \alpha \cdot a_{\pi})$$

Where τ is scaled time and t is unscaled time & gestural activations are a function of scaled time.

$a_{constric}(\tau)$, $a_{\pi}(\tau)$, and $a_{neutral} = (1 - a_{constric})$

Byrd & Saltzman 2003