Task Dynamics
Phonological Units (Gestures) as Dynamical Systems

- Dynamical system is time-invariant, but lawfully produces time-varying state.

- Change in state is continuous, but change in system is discrete.

- System is context independent, but gives rise to context-dependent trajectories of the state variable, as a function of initial conditions.
Task dynamics and gestures

- How do we use dynamical systems to simulate articulator motion?

- Each consonant and vowel can be thought of as a (motor) “task” to achieve a goal for a particular state variable.
  - What might be relevant goals and state variables for consonant and vowel tasks?

- The change over time of the state can be controlled by a dynamical system with a goal and a stiffness.

- The changing state causes changes in the articulators of the vocal tract that can produce those state changes.
  - The relevant articulators for producing a given task are organized as a synergy. Their dynamical equations are linked to one another.
Dynamical system for producing lip closure

• Rule for change:
  • Change in \( x = -kx + C \)

• What is state \((x)\)?
  • Distance between the lips: Lip Aperture (LA)
  • What should the goal state be for lip closures?
  • 0?
  • actually a negative value: the lips compress to form a tight seal.

• What are the values of \( k \) and \( C \)?
  • \( k \) will determine rate at which LA gets to its goal state
  • Goal value is used to set \( C \).
  \[ C = \text{Goal}(LA) \times k \]
Dynamical system for producing lip closure

- Actual Rule for change:

- mass-spring dynamics with critical damping:

\[ \ddot{x} = -\frac{b}{m} \dot{x} - \frac{k}{m} (x - x_0) \]
Compression Goal for lip closure

- Notice that the lips continue to get closer together, even while they are completely closed.

“about”
Cognitive and Physical

\[ \ddot{x} = -\frac{b}{m} \dot{x} - \frac{k}{m} (x - x_0) \]

- The parameter specifications of the system are linguistic.
- Invariant over lifetime of system
- Invariant over contexts.

- \( x \) and its derivatives are continuously changing state variables.

- The are both inextricable parts of a single equation, or law.
TaDA (Task Dynamic Application)

- TaDA synthesizes speech from a list of constriction tasks (also called gestures) to be produced by dynamical systems.
- Each task is specified for:
  - Identity of Task variable (state)
  - Interval of time during which the dynamics of this task controls the task variable
  - Goal value
  - Stiffness
  - Articulators that form the synergy for that task, and their average relative contributions or weights (big numbers mean the articulator is “heavy” and contributes less).
  - Blending parameters ($\alpha$ and $\beta$ for when more than one Task controls the same variable at the same time (overlap, like we saw with the thermostat with leaky insulation)).

<table>
<thead>
<tr>
<th>Task</th>
<th>Beg Fr</th>
<th>End Fr</th>
<th>Goal</th>
<th>k</th>
<th>Articulator Weights</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>'LA'</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>-2</td>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>
TaDA Synthesis of speech

- Simulate motion of articulators when producing a particular word using Task Dynamics.
- Calculate time-varying vocal tract shape from positions of articulators using vocal tract model (CASY).
- Calculate time-varying filter action of vocal tract.
- Create simulated laryngeal vibration (and any voiceless sources) and filter through the time-varying vocal tract shape.
CASY: Vocal Tract Model

Articulator Positions

Tract Outline

Tube Shape

Formants:
515
1708
2569
3292
4233

Frequency in Hz

Magnitude in dB
TADA synthesis: synergies

- The articulators available to the Tasks in TADA are the CASY articulators.
- TADA incorporates knowledge of the relation between articulator positions and the values of the task variables.
- As the task variable changes over time (due to the dynamical system), a model of the synergy calculates the change in CASY articulator positions.
- Given articulator positions, CASY calculates vocal tract shapes, areas and formants.
TADA Synthesizer: CASY Articulators for LA task
TaDA: “bin ban”

- The line specifying lip closure task is the same for both [b]s.
  - ’LA’ 0 0 9 0 -2 8 1 \(JA = 8\), \(UH = 5\), \(LH = 1\) 100 0.01
  - ’LA’ 0 34 43 0 -2 8 1 \(JA = 8\), \(UH = 5\), \(LH = 1\) 100 0.01
- But due to the low jaw in “ban”, the upper lip lowers more than it does in ”bin”, and the lower lip raises more.
- This emerges automatically from synergy model.
TaDA: “bin ban”

“bin”

“ban”

Time (ms)
Synthesis of syllables

- We need to know what the tasks are for vowels and other consonants.
- We need to know how to arrange them in time.
- Let’s try to synthesize the word “bob”.
- Need to find the task for vowels: Constriction Location and Degree of tongue body.
## Tasks

<table>
<thead>
<tr>
<th>Segments</th>
<th>Task Variables</th>
<th>Articulators</th>
</tr>
</thead>
<tbody>
<tr>
<td>p b m</td>
<td>Lip Aperture (LA)</td>
<td>Upper Lip (uh) Lower Lip (lh) Jaw (ja)</td>
</tr>
<tr>
<td>t d n s z sh zh</td>
<td>Tongue Tip: Constriction Degree (TTCD) Constriction Location (TTCL)</td>
<td>Tongue Tip (ta, tl) Tongue Body (cl, ca) Jaw (ja)</td>
</tr>
<tr>
<td>k g n x Vowels</td>
<td>Tongue Body: Constriction Degree (TBCD) Constriction Location (TBCL)</td>
<td>Tongue Body (cl, ca) Jaw (ja)</td>
</tr>
<tr>
<td>m n n x</td>
<td>Velic Aperture (VEL)</td>
<td>Velum (na)</td>
</tr>
<tr>
<td>p t k s sh</td>
<td>Glottal Aperture (GLO)</td>
<td>Glottal Width (gw)</td>
</tr>
<tr>
<td>uw uh ow ao</td>
<td>Protrusion (PRO)</td>
<td>Lip Protrusion (lp)</td>
</tr>
</tbody>
</table>
TBCL and TBCD for Vowels

TBCL=95
TBCD=4

TBCL=180
TBCD=11

IY

AA
CASY Tongue Articulators for TB tasks
Gestural Score: Putting gestures together in time

- Fundamental Principle: CV co-production
- The gestures for the syllable-initial C and the V begin at the same time.
Gestural Score: CVC

- The V gesture continues longer than initial C.
- The gesture for the final C begins after the initial C ends with delay equal to about the duration of a C task.
Gestural Score

“two back”

AUGMO

400 500 600 700 800 900 1000 1100 1200
Gestural Score

“had made him lose”
Gestural Score for “bob”

```
% /b/
'LA' 0 0 9 0 -2 8 1 JA=8, UH=5, LH=1 100 0.01
'LA' 0 9 13 0 11 8 1 JA=8, UH=5, LH=1 1 1
'VEL' 0 0 9 0 -0.1 8 1 NA=1 0 0

% /aa/
'TBCD' 0 0 30 0 11 4 1 JA=1, CL=1, CA=1 1 1
'TBCL' 0 0 30 0 170 4 1 JA=1, CL=1, CA=1 1 1

% /b/
'LA' 0 25 34 0 -2 8 1 JA=8, UH=5, LH=1 100 0.01
'LA' 0 34 37 0 11 8 1 JA=8, UH=5, LH=1 1 1
'VEL' 0 25 34 0 -0.1 8 1 NA=1 0 0
```
“bob” to “mob”

- Change initial VEL goal from -0.1 to 0.2

```
%/m/
'LA' 0 0 9 0 -2 8 1 JA=8, UH=5, LH=1 100 0.01
'LA' 0 9 13 0 11 8 1 JA=8, UH=5, LH=1 1 1
'VEL' 0 0 9 0 0.2 8 1 NA=1 0 0

%/aa/
'TBCD' 0 0 30 0 11 4 1 JA=1, CL=1, CA=1 1 1
'TBCL' 0 0 30 0 170 4 1 JA=1, CL=1, CA=1 1 1

%/b/
'LA' 0 25 34 0 -2 8 1 JA=8, UH=5, LH=1 100 0.01
'LA' 0 34 37 0 11 8 1 JA=8, UH=5, LH=1 1 1
'VEL' 0 25 34 0 -0.1 8 1 NA=1 0 0
```
Glottal Aperture Task for voiceless stops and fricatives

- How do we change “bob” to “pop”?
- Goal for GLO = 0.3 for voiceless stops and fricatives
- Default during speech: GLO=0.0 (will produce voicing)
“pop”

All the tasks remain the same, we just add a new one.
Vowel Tasks

- IY: TBCL=95, TBCD=5
- IH: TBCL=95, TBCD=8
- EH: TBCL=95, TBCD=11.5
- AE: TBCL=95, TBCD=17
- UW: TBCL=125, TBCD=4
- OW: TBCL=150, TBCD=5
- AA: TBCL=170, TBCD=11
Tongue Tip Tasks

for /t,d,n/
Articulators for TT Tasks

And also Jaw Angle (JA), Tongue Center (CL, CA)
“beat”

%/b/
'LA' 0 0 9 0 -2 8 1 JA=8, UH=5, LH=1 100 0.01
'LA' 0 9 13 0 11 8 1 JA=8, UH=5, LH=1 1 1
'VEL' 0 0 9 0 -0.1 8 1 NA=1 0 0

%/iy/
'TBCD' 0 0 30 0 4 4 1 JA=1, CL=1, CA=1 1 1
'TBCL' 0 0 30 0 95 4 1 JA=1, CL=1, CA=1 1 1

%/t/
'TTCD' 0 25 34 0 -2 8 1 JA=32, CL=32, CA=32, TL=1, TA=1 100 0.01
'TTCL' 0 25 34 0 56 8 1 JA=32, CL=32, CA=32, TL=1, TA=1 1 1
'TTCL' 0 34 39 0 24 8 1 JA=512, CL=512, CA=512, TL=1, TA=1 1 1
'TTCD' 0 34 39 0 11 8 1 JA=512, CL=512, CA=512, TL=1, TA=1 1 1
'VEL' 0 25 34 0 -0.1 8 1 NA=1 0 0
'GLO' 0 28 39 0 0.4 16 1 GW=1 0 0
Tasks for velar stops?
### Tasks for stop consonants

<table>
<thead>
<tr>
<th>Lip (Labial)</th>
<th>T Tip (Alveolar)</th>
<th>T Body (Velar)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CLO</strong></td>
<td><strong>REL</strong></td>
<td><strong>CLO</strong></td>
</tr>
<tr>
<td>LA=-2</td>
<td>LA=11</td>
<td>TTCL=56</td>
</tr>
<tr>
<td>b VEL=-.1</td>
<td>d VEL=-.1</td>
<td>g VEL=-.1</td>
</tr>
<tr>
<td>p VEL=-.1</td>
<td>t VEL=-.1</td>
<td>k VEL=-.1</td>
</tr>
<tr>
<td>m VEL=.2</td>
<td>n VEL=.2</td>
<td>nx VEL=.2</td>
</tr>
<tr>
<td>VEL=.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VEL=.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VEL=.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Linguistics 285 (USC Linguistics)
Lecture 15: Speech Synthesis and Task Dynamics
March 11, 2014 20 / 21
% /b/
% clo
'LA' ... -2 8 1 JA=8,UH=5,LH=1 100 0.01
% rel
% 'LA' ... 9 13 0 11 8 1 JA=8,UH=5,LH=1 1 1

% /d/
% clo
'TTCL' ... 56 8 1 JA=32,CL=32,CA=32,TL=1,TA=1 1 1
'TTCD' ... -2 8 1 JA=32,CL=32,CA=32,TL=1,TA=1 100 0.01
% rel
'TTCL' ... 24 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'TTCD' ... 11 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1

% /g/
% clo
'TBCL' ... 110 8 1 JA=10,CL=1,CA=1 10 0.1
'TBCD' ... -2 8 1 JA=100,CL=1,CA=1 100 0.01
% rel
'TBCD' ... 6 8 1 JA=10,CL=1,CA=1 1 1

Stops: Oral Constriction Gestures
Glottal and Velic gestures for stops and fricatives

% Velic closure for stops and fricatives

'VEL' ... -0.1 8 1 NA=1 0

% Velic opening for nasals

'VEL' ... 0.2 8 1 NA=1 1 1

% Glottal opening for voiceless stops and fricatives

'GLO' ... 0.3 16 1 GW=1 0
Fricatives: Oral Constriction Tasks

θ

z

ʃ
# Fricatives: Oral Constriction Tasks

<table>
<thead>
<tr>
<th></th>
<th>θ</th>
<th>s</th>
<th>∫</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TT</strong></td>
<td>clo</td>
<td>rel</td>
<td>clo</td>
</tr>
<tr>
<td></td>
<td>TTCL=40 TTCD=1.2</td>
<td>TTCL=24 TTCD=11</td>
<td>TTCL=56 TTCD=1.2</td>
</tr>
<tr>
<td><strong>TB</strong></td>
<td></td>
<td></td>
<td>TTCL=110 TBCD=10</td>
</tr>
</tbody>
</table>
Fricatives: examples

\( \theta \) \hspace{1cm} s \hspace{1cm} \int
Fricatives:
  Oral
  Constriction
  Gestures

% /th/
% clo
'TTCL' ... 40 10 1 JA=32,CL=32,CA=32,TL=1,TA=1 1 1
'TTCD' ... 1.2 10 1 JA=32,CL=32,CA=32,TL=1,TA=1 10 0.1

% rel
'TTCL' ... 24 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'TTCD' ... 11 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1

% /s/
% clo
'TTCL' ... 56 10 1 JA=640,CL=32,CA=32,TL=1,TA=1 1 1
'TTCD' ... 1.2 10 1 JA=640,CL=32,CA=32,TL=1,TA=1 10 0.1
'TBCL' ... 110 8 1 JA=10,CL=1,CA=1 100 0.1
'TBCD' ... 10 8 1 JA=10,CL=1,CA=1 100 0.1

% rel
'TTCL' ... 24 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'TTCD' ... 11 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1

% /sh/
% clo
'TTCL' ... 60 10 1 JA=640,CL=32,CA=32,TL=1,TA=1 1 1
'TTCD' ... 1.2 10 1 JA=640,CL=32,CA=32,TL=1,TA=1 10 0.1
'TBCL' ... 95 8 1 JA=10,CL=1,CA=1 100 0.01
'TBCD' ... 8 8 1 JA=10,CL=1,CA=1 100 0.01

% rel
'TTCL' ... 40 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'TTCD' ... 11 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
Glides: Oral Constriction Gestures

% /w/
%clo
'LA' ... 181 JA=8, UH=5, LH=1 1 1
'TBCL' ... 12581 JA=10, CL=1, CA=1 10 0.1
'TBCD' ... 281 JA=10, CL=1, CA=1 100 0.01

%rel
'LA' ... 1181 JA=8, UH=5, LH=1 1 1

% /j/
'TBCL' ... 9581 JA=10, CL=1, CA=1 100 0.01
'TBCD' ... 281 JA=10, CL=1, CA=1 100 0.01