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?

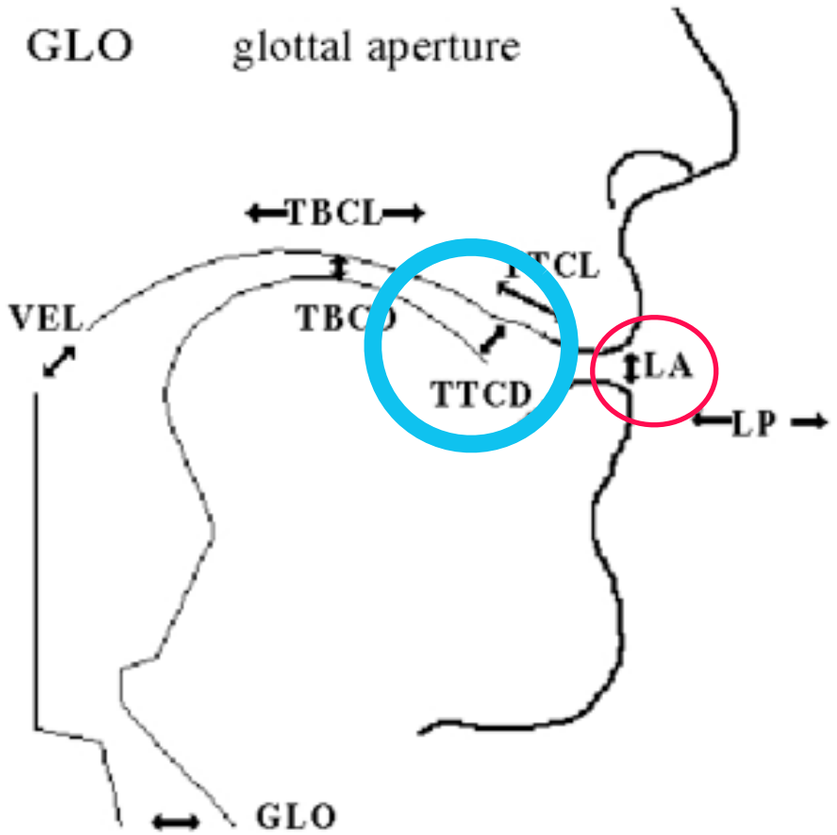
- The change over time of the task variable can be controlled by a simple dynamical system with a goal and a stiffness.

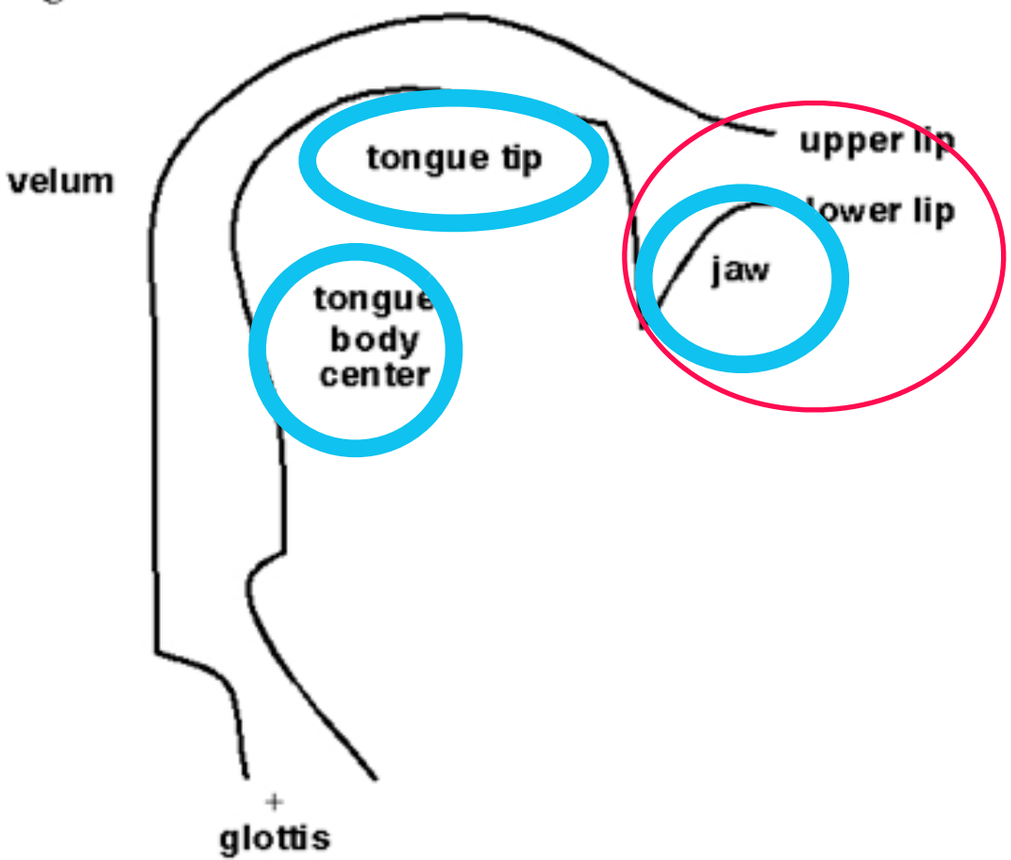
- The changing task state causes changes in the articulators of the vocal tract that can produce those state changes.

  - The relevant articulators for producing a given task are organized as a synergy. Their dynamical equations are linked to one another.

Gesture task variables

	Task	Articulators
LP	lip protrusion	upper & lower lips, jaw
LA	lip aperture	upper & lower lips, jaw
TTCL	tongue tip constrict location	tongue tip, tongue body, jaw
TTCD	tongue tip constrict degree	tongue tip, tongue body, jaw
TBCL	tongue body constrict location	tongue body, jaw
TBCD	tongue body constrict degree	tongue body, jaw
VEL	velic aperture	velum
GLO	glottal aperture	glottis



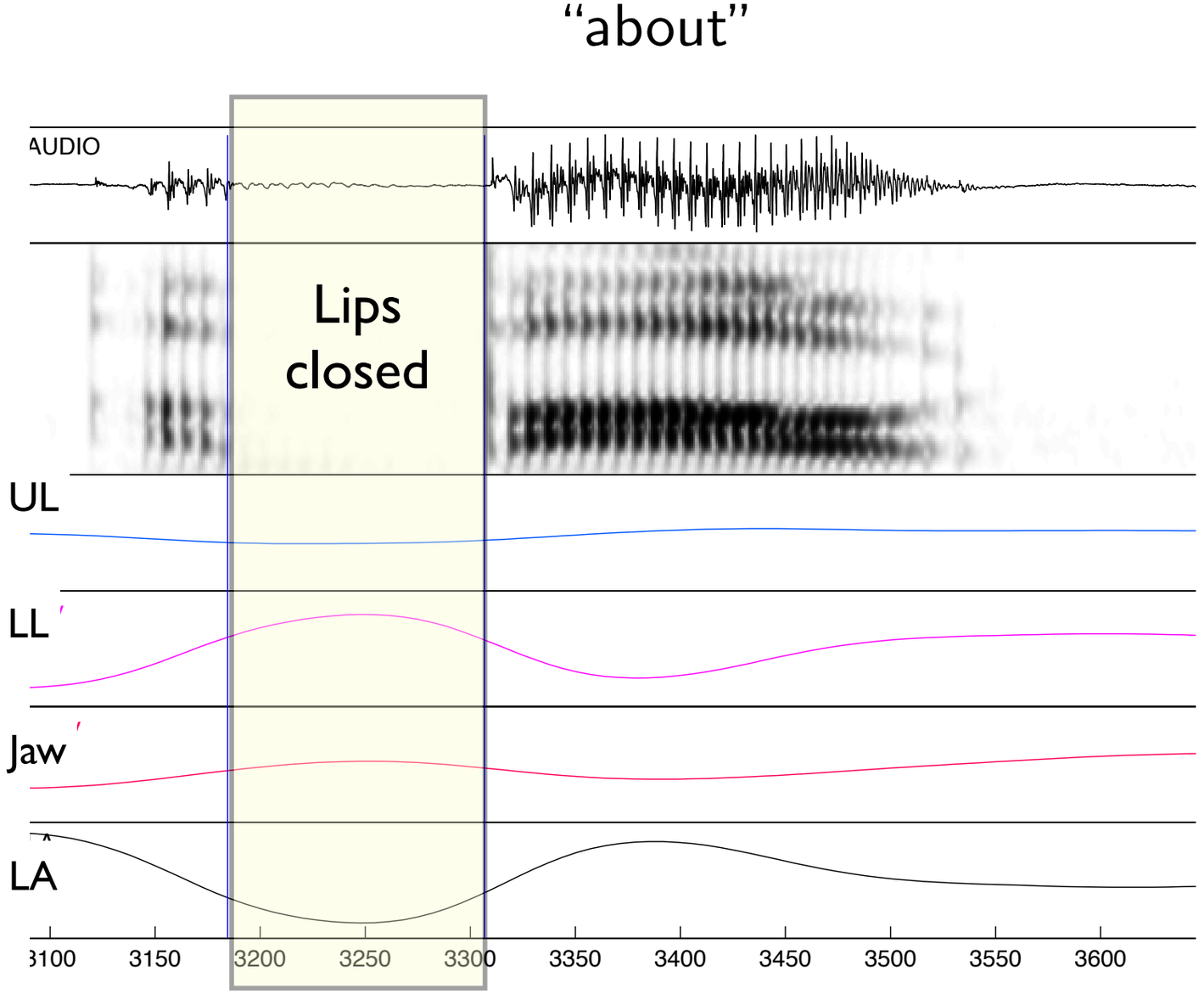


Dynamical system for producing lip closure

- 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= Goal(LA) * k
-

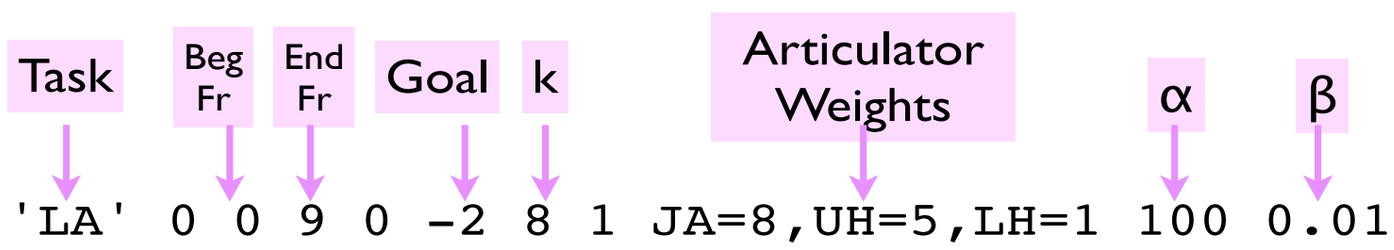
Compression Goal for lip closure

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

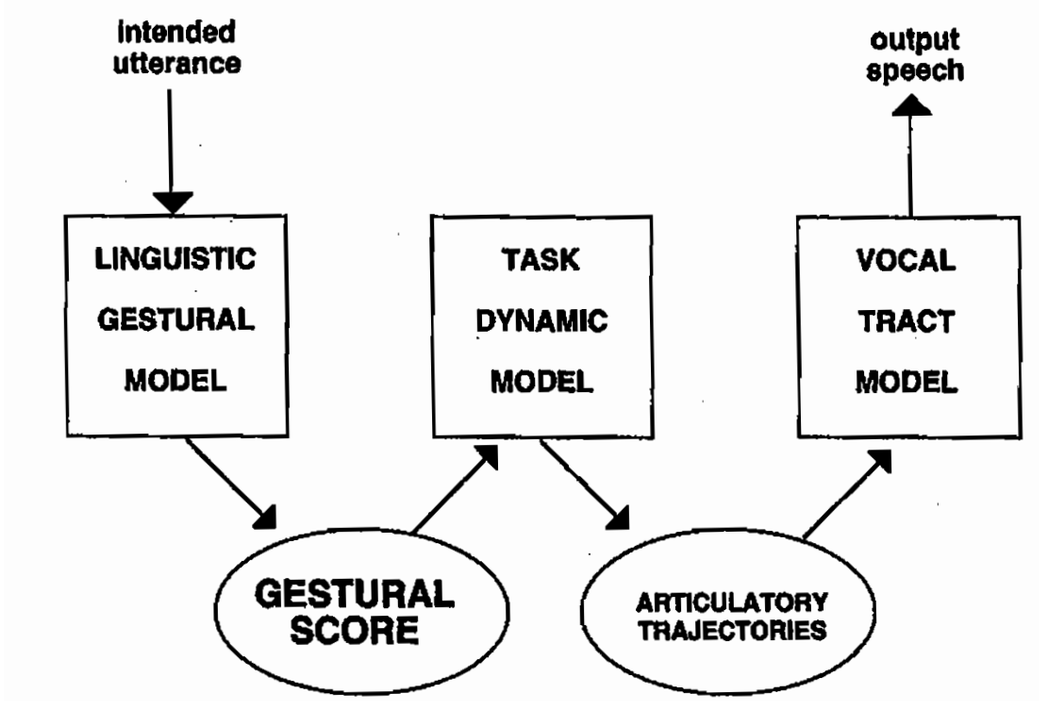


TaDA (Task Dynamic Application)

- TaDA synthesizes speech from a list of constriction tasks (also called *gestures*) to be produced by dynamical systems. This list is a gestural score.
- Each task is specified for:
  - Identity of Task variable (state)
  - Interval of time during which the dynamics of this task controls the task variable state.
  - 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 (α and β for when more than one Task controls the same variable at the same time (overlap, like we saw with the thermostat with leaky insulation).

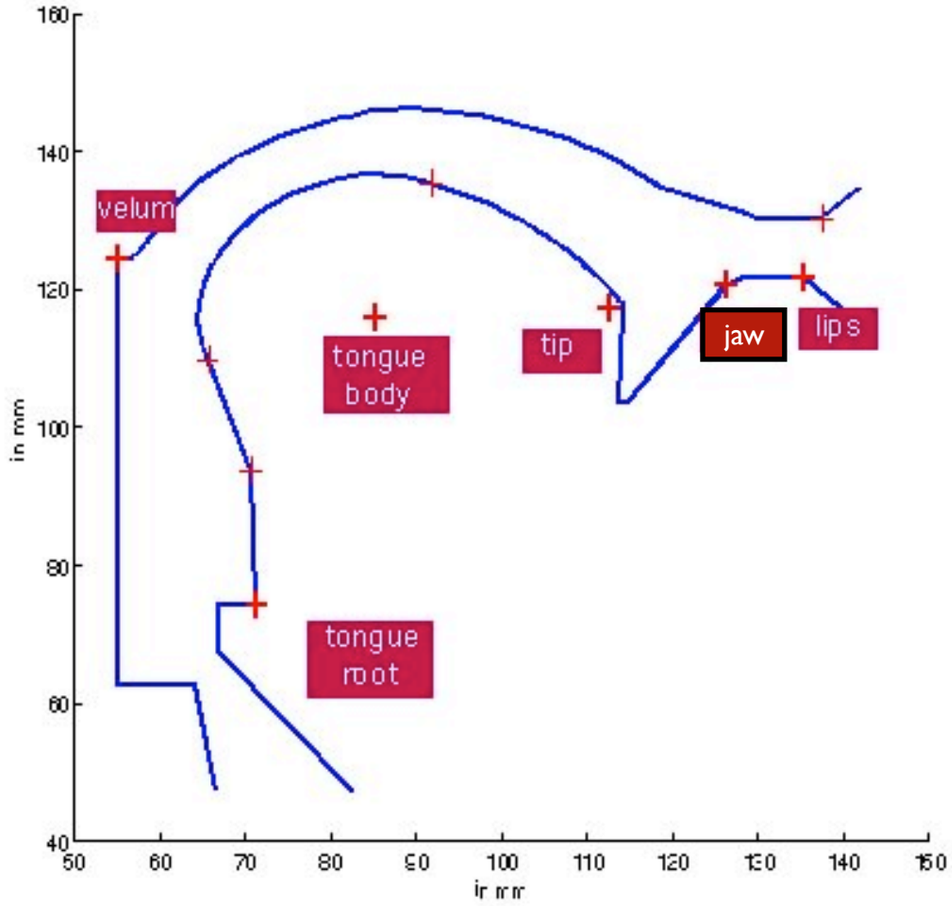


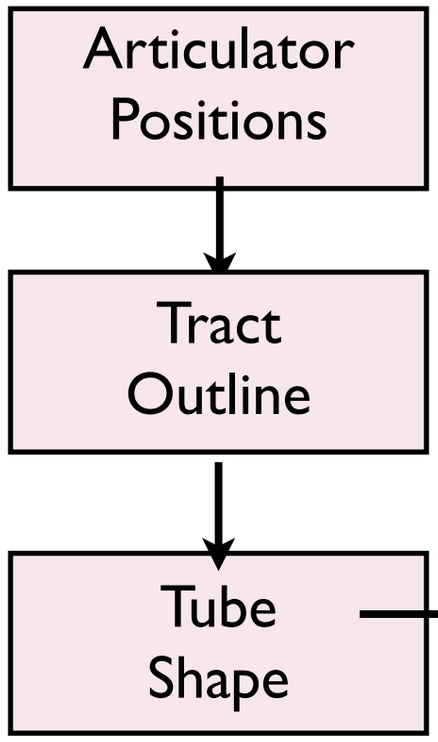
TaDA Synthesis of speech



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

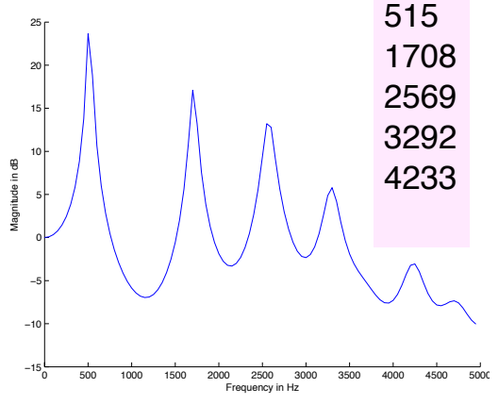
CASY: Vocal Tract Model





Formants

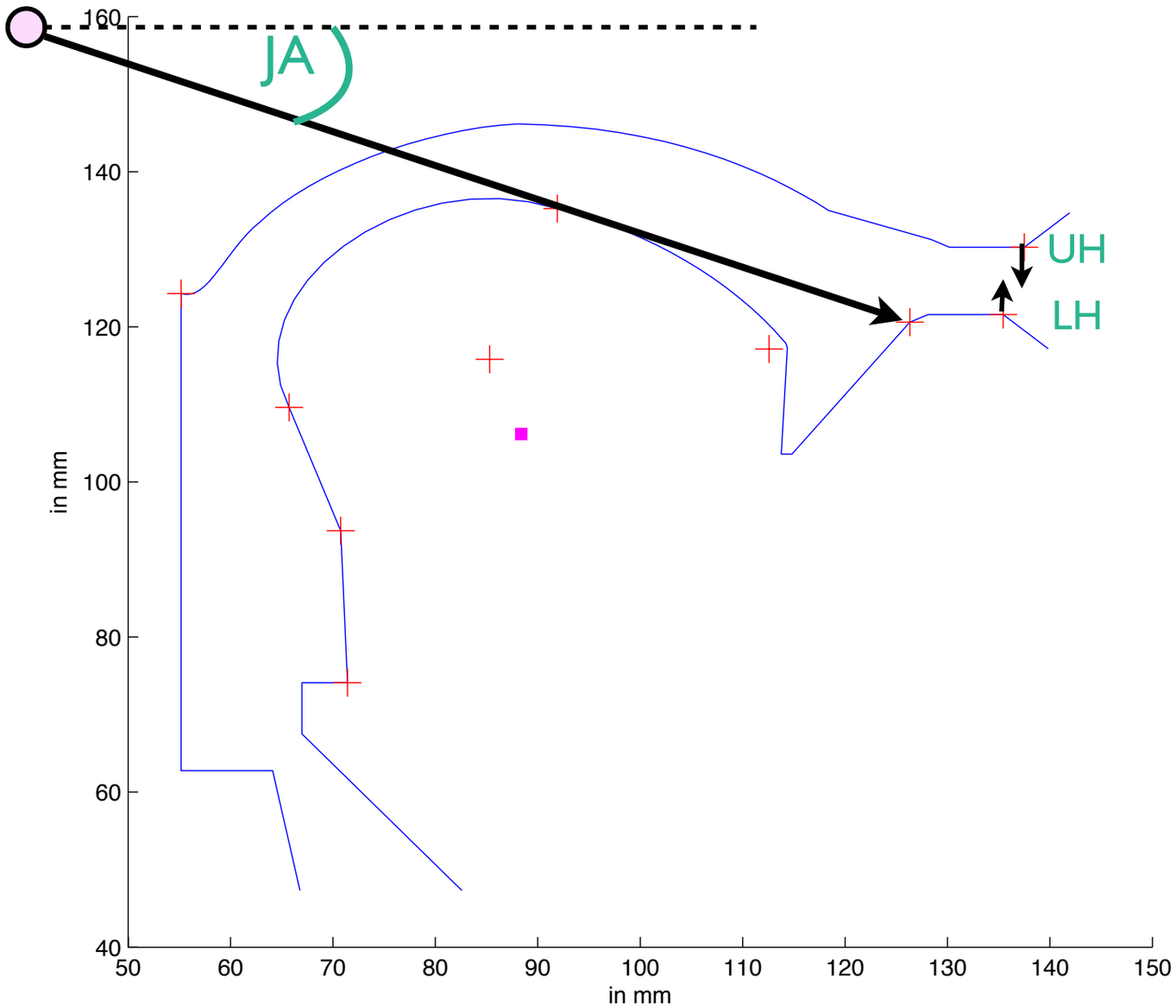
515
1708
2569
3292
4233



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 variable state. (TV state)
- As the TV state 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



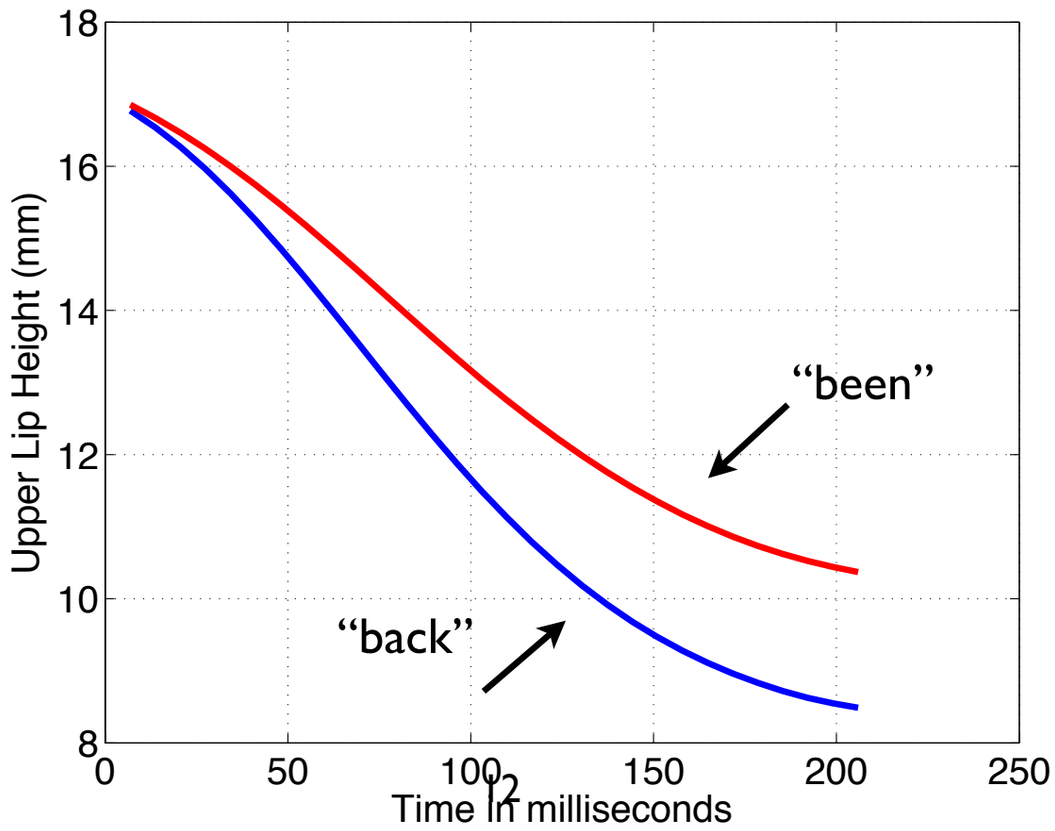
Lip Task performance in different contexts

- The relative contribution of the articulators in the synergy may differ when the task is produced in different contexts in which one of the articulators may be required for some other task.

- For example, lip closure in "back" vs. "been".

- Jaw is recruited to be low in "back" because of the low vowel (/ae/) and high in "been" because of high vowel (/ɪ/),

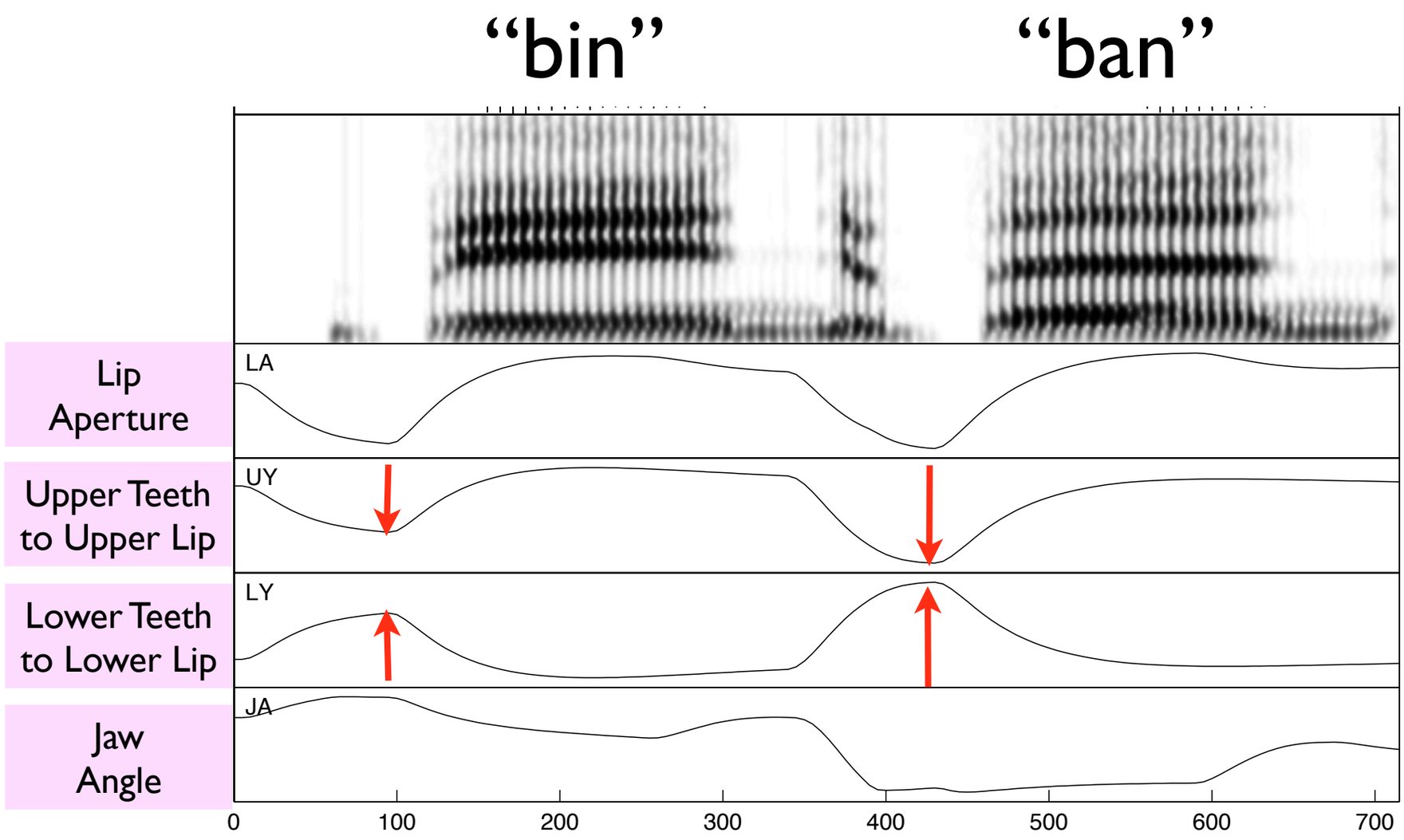
- More upper lip lowering emerges in "back" than in "been".



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"
Lip Aperture
Upper Teeth to Upper Lip
Lower Teeth to Lower Lip
Jaw Angle
LA
UY
LY
JA
0
100
200
300
400
500
600
700
Time (ms)

Synthesis of syllables

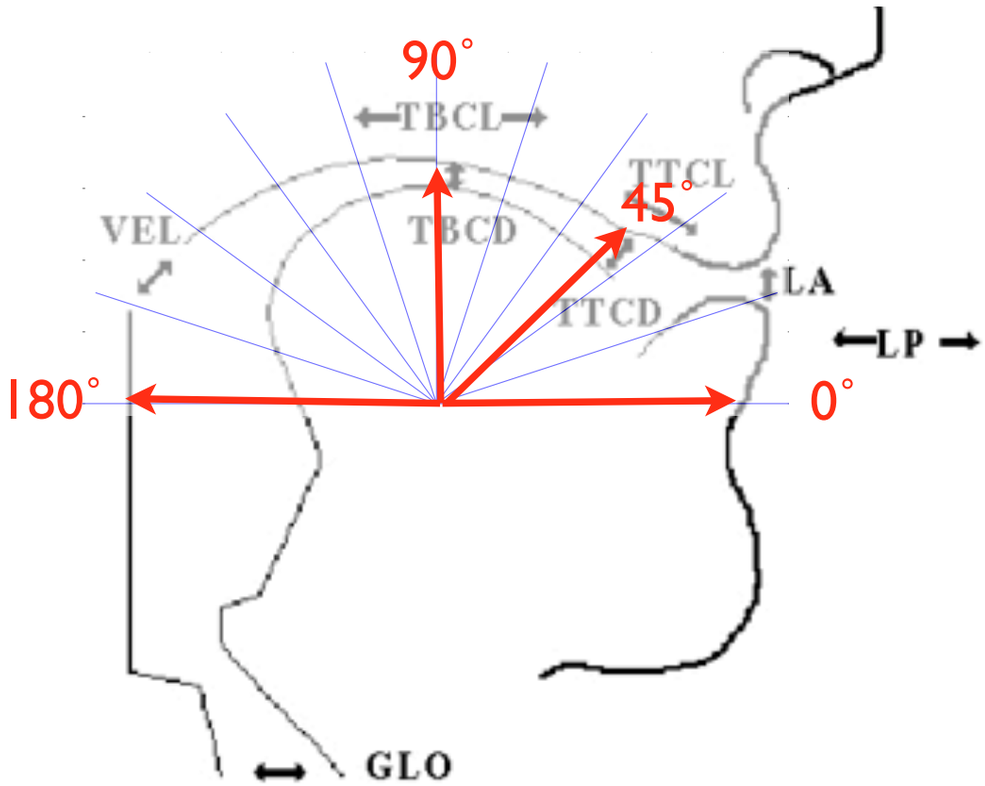
- We need to know what the tasks are for vowels and other consonants

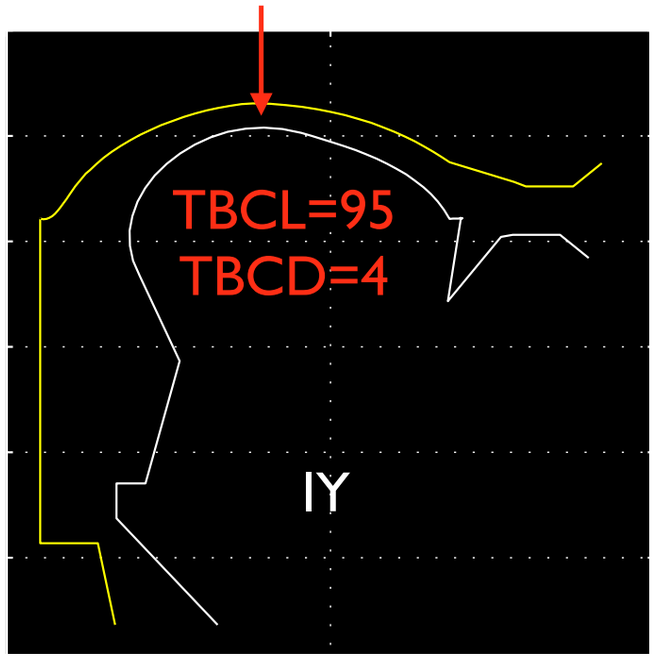
- We need to know how to arrange them in time, how long they are active and what their stiffnesses are.

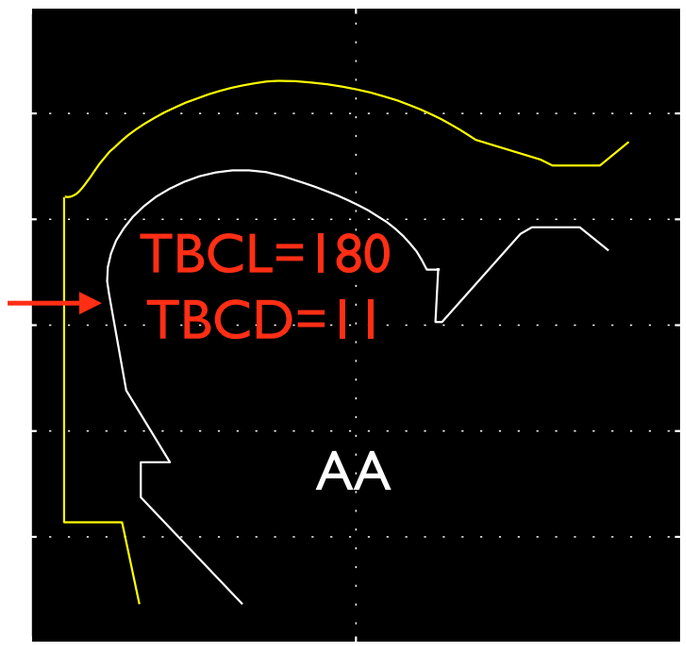
- Need to find the task for vowels:
  Constriction Location and Degree of tongue body.

- Let's try to synthesize the word "bob".

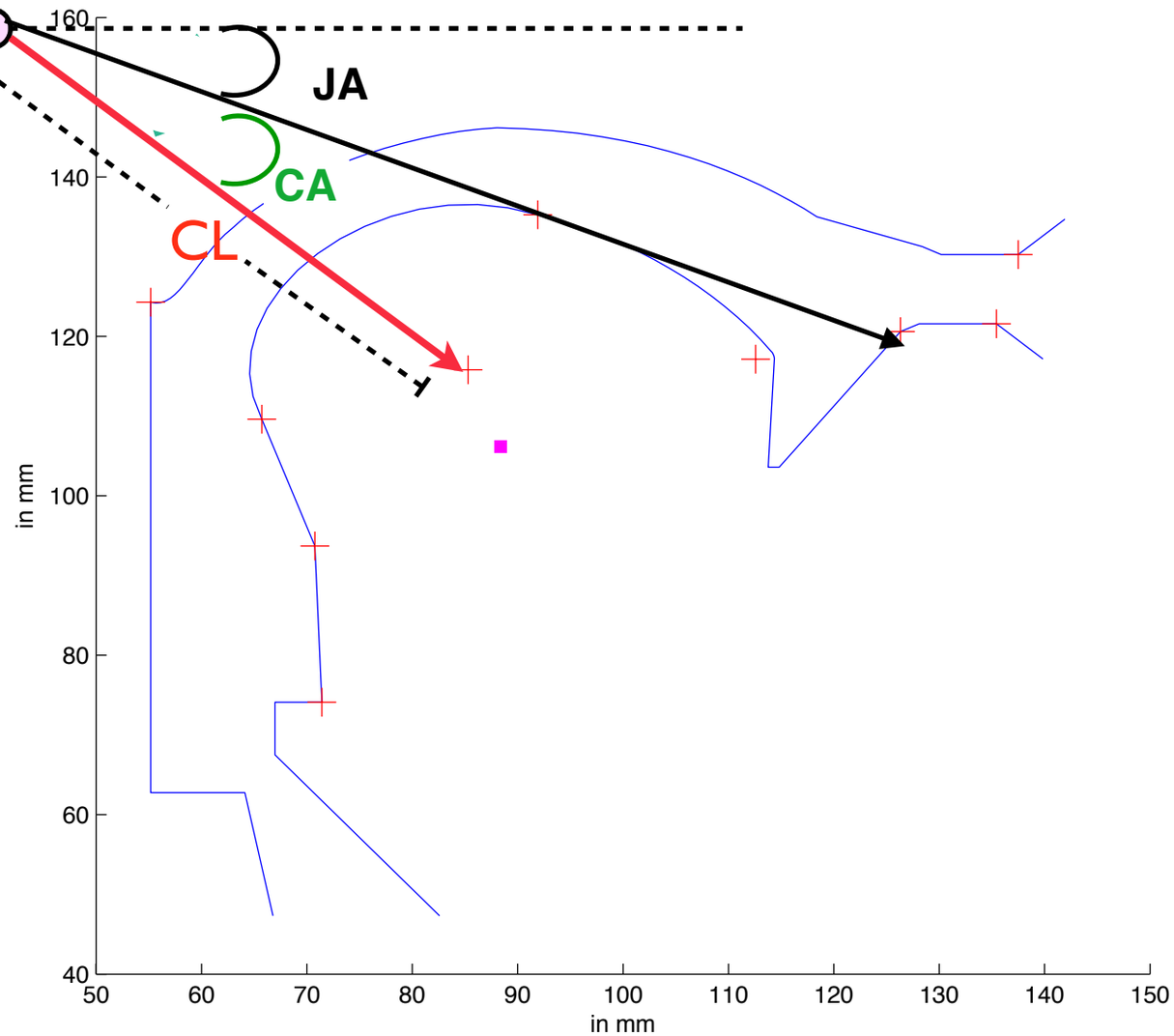
TBCL and TBCD for Vowels







CASY Tongue Articulators for TB tasks

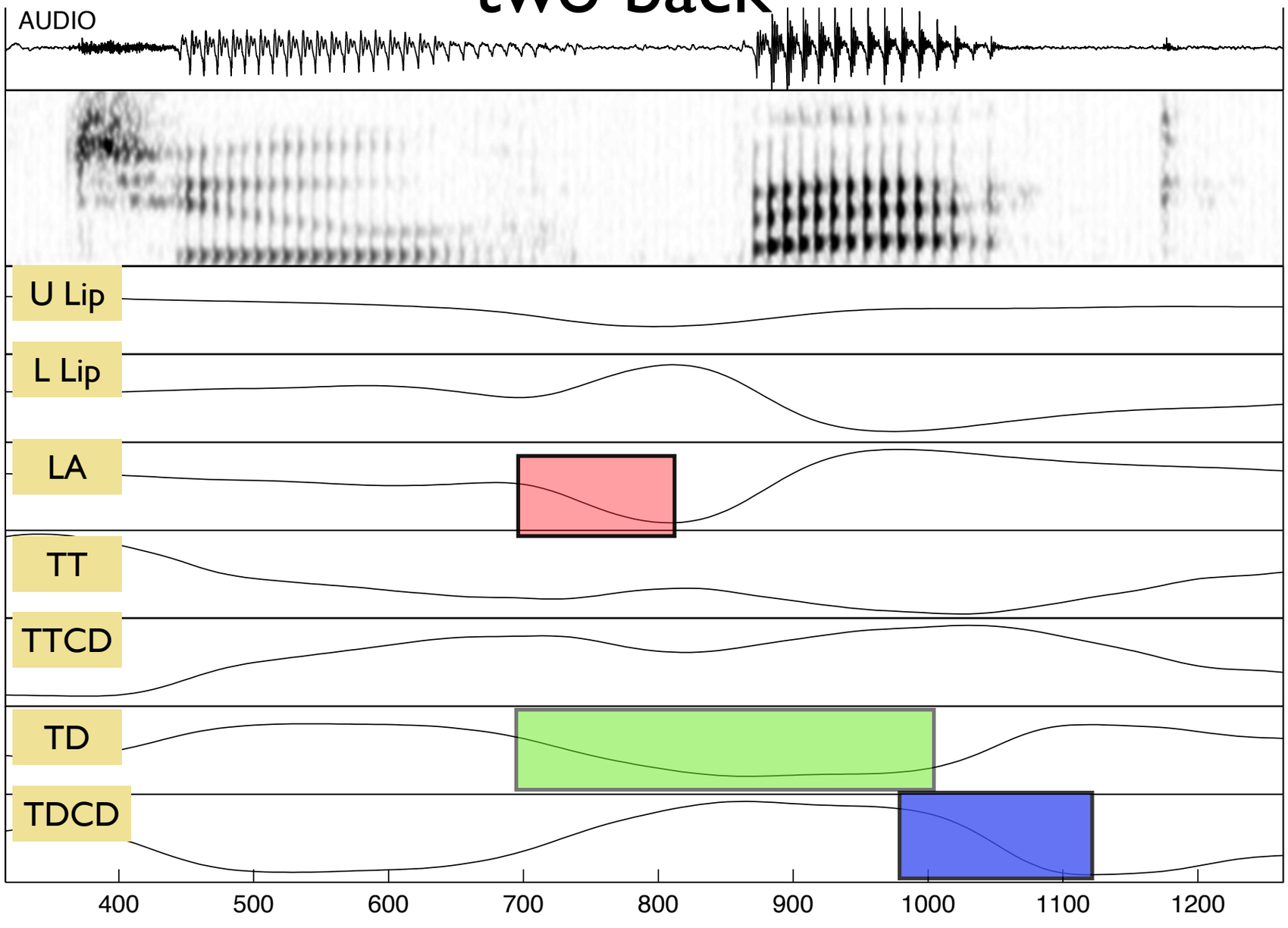


Gestural Score: Putting gestures together in time

- Fundamental Principle: CV co-production

- The gestures for the syllable-initial C and the V begin at the same time.

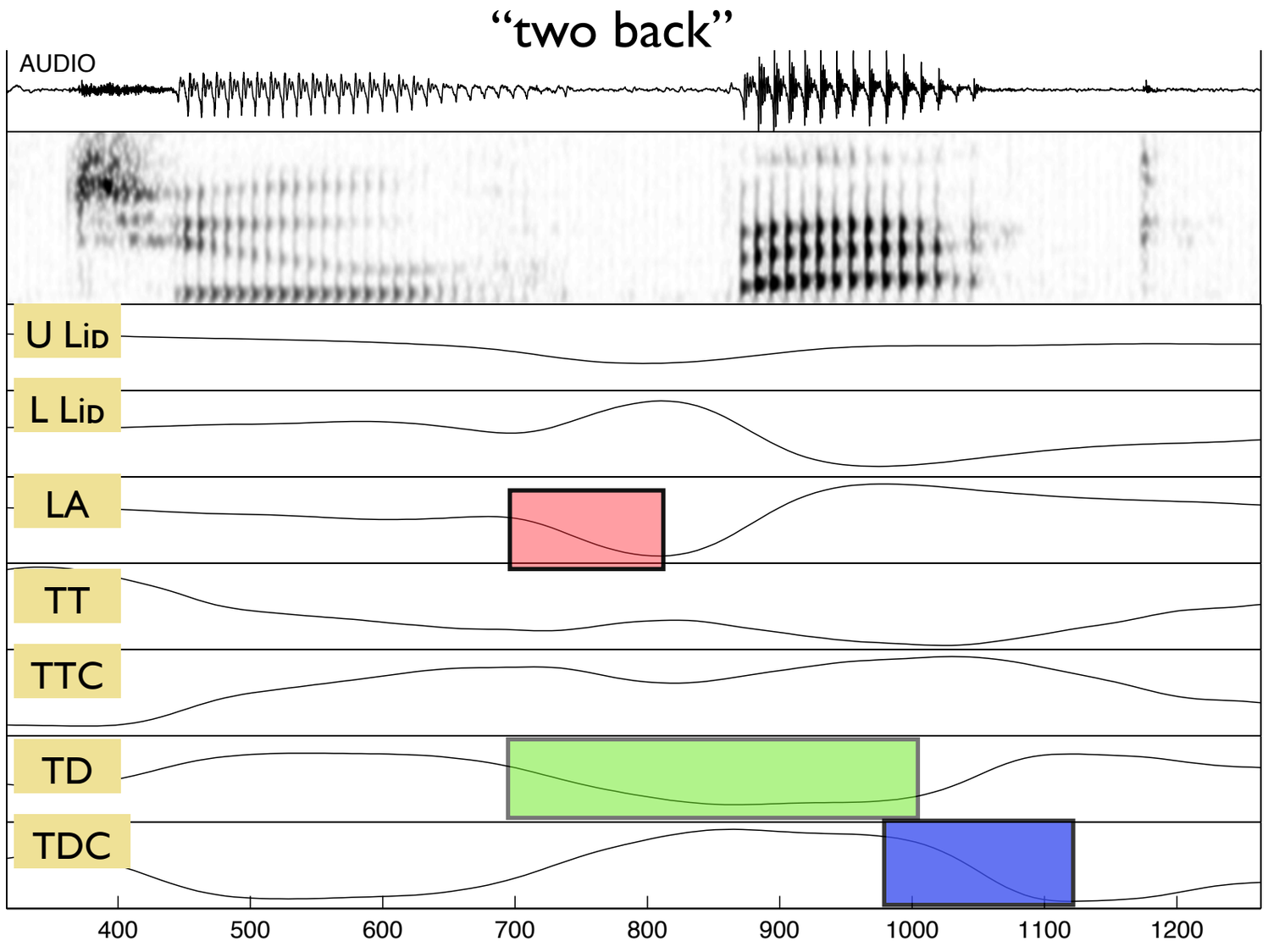
"two back"



Note that the movement into the consonant is faster than the movement into vowel

Gestural Score: CVC

- The V gesture is about twice as long as the initial C.
- The gesture for the final C begins near the end of the V task.
- Stiffness for consonant gestures is higher than that of vowel gestures (Consonants = 8 Hz, Vowels = 4 Hz)



Additional (non-contrastive) consonant gestures

We have focussed our discussion and homework exercises on the formation of contrastive oral constrictions in the vocal tract.

However, additional (largely) non-contrastive gestures are required to allow these constrictions to have robust audible acoustic consequences.

(1) Release Gestures

—The dynamical system that controls the constriction formation is active for some time interval and results in the constriction being achieved.

—When that interval is over, the articulators are free to do other things, but to ensure that the articulators move away from the constricted posture rapidly enough for consonants, a release gesture is required to regulate the increase in aperture of the task variable.

—Release gestures immediately follow the constriction gesture in time.

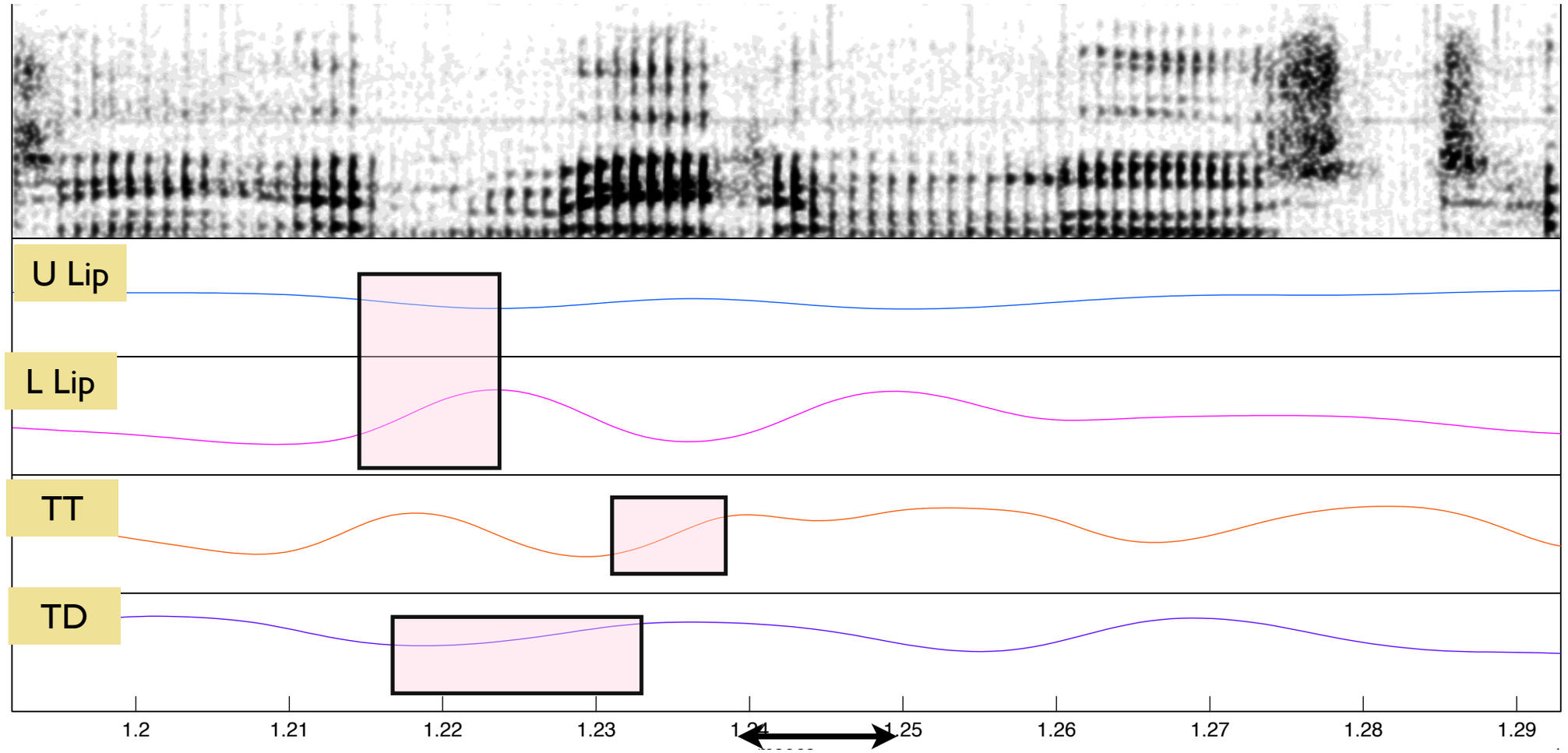
(2) Velic Closure

For stops and fricatives, the velum must be firmly closed.

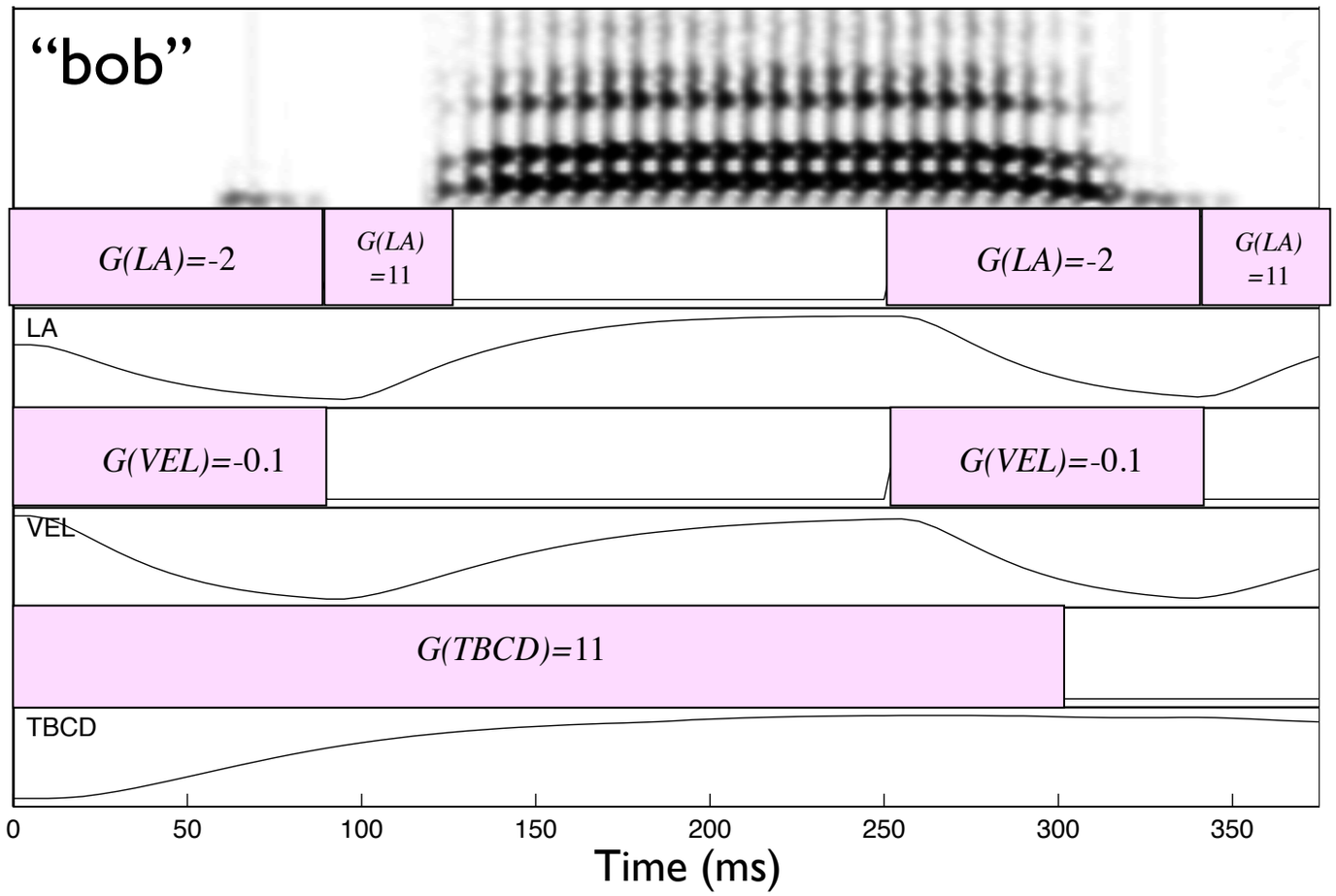
A velic closure gesture ensures this.

Gestural Score

"had made him lose"



Gestural Score for “bob”



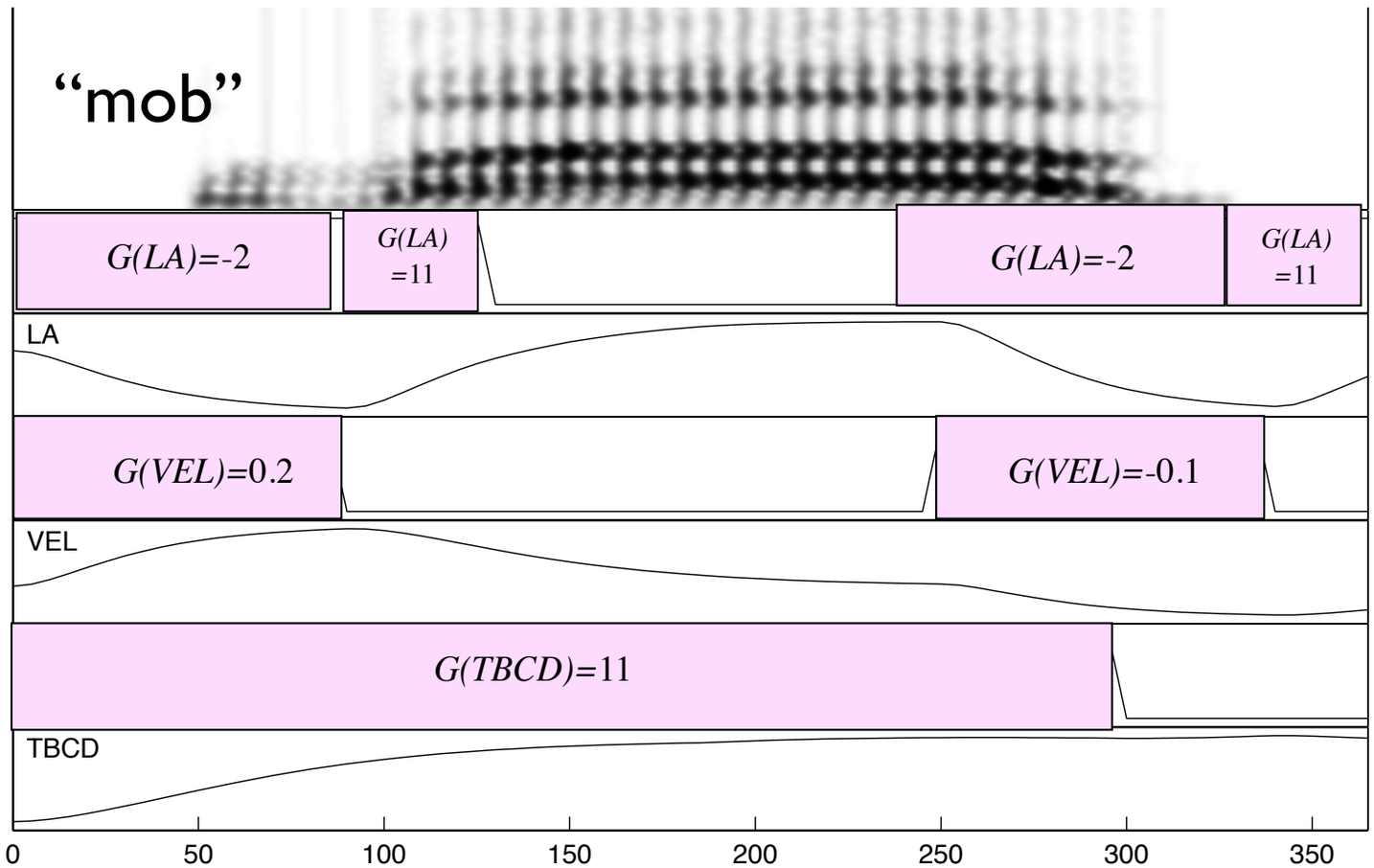

```
%/b/
'LA' 0 0 9 0 -2 8 1 JA=8,UH=5,LH=1 100 0.01
'LA' 0 9 13 0 11 8 1 JA=8,UH=5,LH=1 1 1
'VEL' 0 0 9 0 -0.1 8 1 NA=1 0 0

%/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
```