

Task Dynamics Application (TADA) for articulatory speech synthesis

Jaekoo Kang

Speech-Language-Hearing Sciences
CUNY Graduate Center
Haskins Laboratories

2020-10-25

Plans

Tutorial 1 (Oct. 19)

Matlab and TADA basics

- Part 1. Matlab basics
- Part 2. Intro to TADA

Tutorial 2 (Oct. 26)

Synthesis using TADA

- Vowel gestures & timing.
- Consonant gestures & timing.

Today!

Q & A

- Matlab
- TADA
 - Issues running TADA?
 - Synthesizing a word stimulus?

Overview

- Review on Matlab and TADA
- Vowel gestures
- Consonant gestures
- Gestural overlap
- Gestural timing

Matlab

- Variables
- Operations
- Loops
- Plotting
- File I/O

The image shows the Matlab interface with three main panes:

- Command Window:** Displays command-line input and output.
- Workspace:** Shows the current variables and their values.
- Editor:** Displays Matlab code for loops, plotting, sound file operations, and a script for audio processing.

Command Window (Variables):

```
>> a = 3;
>> b = 'hello!';
>> c = [1, 0.8];
>> d = [1 2; 3 4];
>> e = {0.2, 'k'};
fx >>
```

Workspace:

Name	Value
a	3
b	'hello!'
c	[1,0.8000]
d	[1,2;3,4]
e	1x2 cell

Editor (Loops):

```
1 for i = 1:10
2     if i < 5
3         a(i) = i^2;
4     end
5 end
6
```

Editor (Plotting):

```
1 x = rand(1,100);
2 plot(x)
3 plot(x, 'r')
4 plot(x, 'ro')
5 hist(x)
6 boxplot(x)
```

Editor (File I/O):

```
1 % Loading a sound file
2 [y, sr] = audioread('sound.wav');
3
4 % Playing a sound file
5 soundsc(y, sr)
```

TADA

- TADA is a task-dynamics application for simulating gestural structure of speech and generating acoustic output.
- Gestures follow rules of dynamical systems.
- A set of invariant parameters generates variable movement over time.

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

“How fast and stable a movement is toward a goal”

k

b

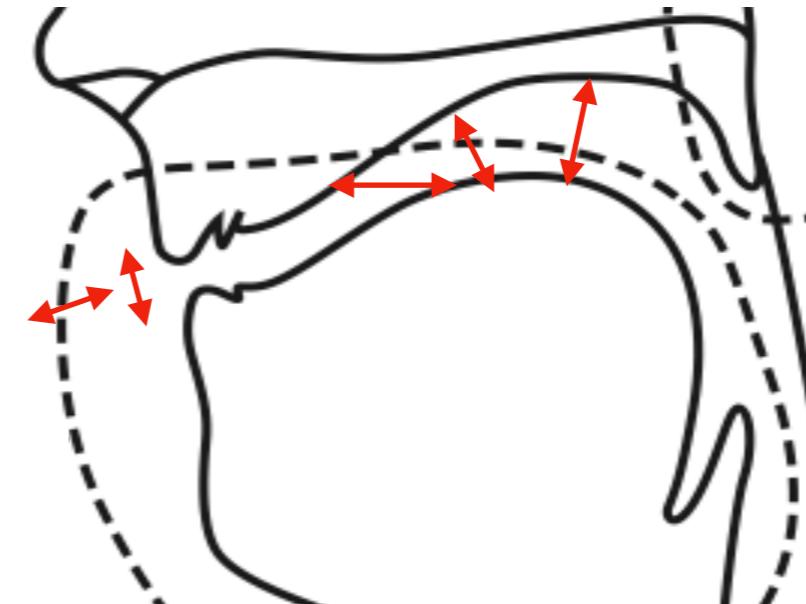
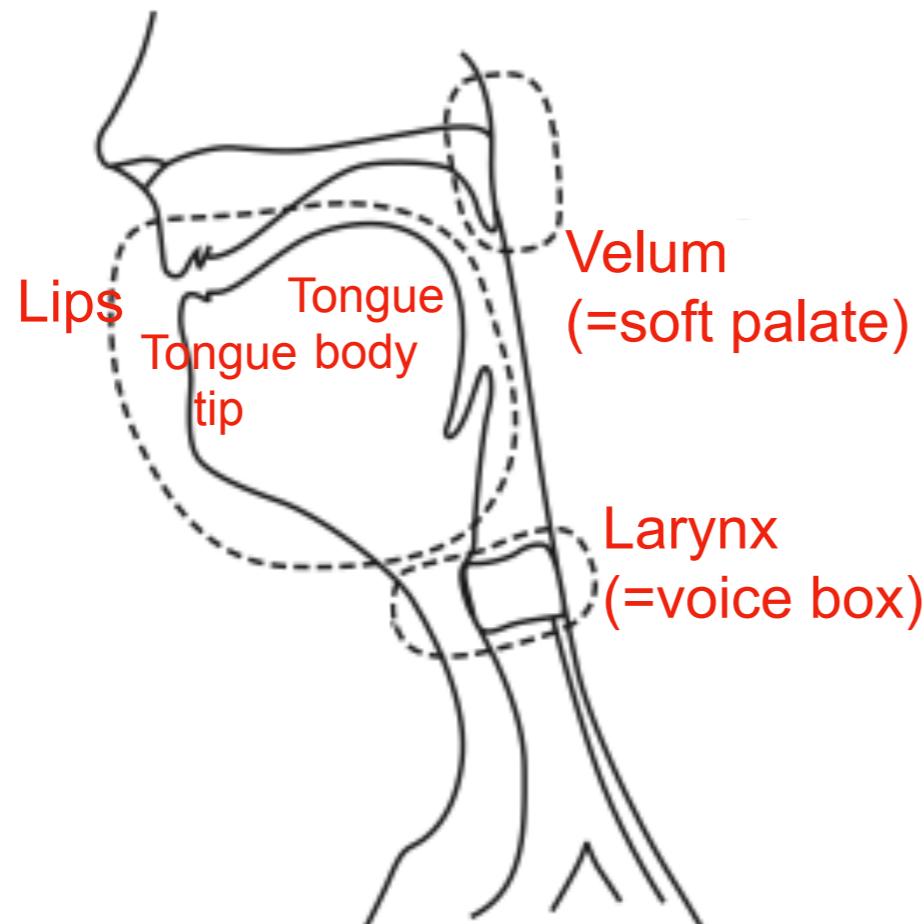
x_0

Speech organs and tasks

- Five major articulators produce gestures of speech.
- Gestures coordinate to achieve a vocal-tract constriction.
 - Constriction Location (CL) and Constriction Degree (CD).

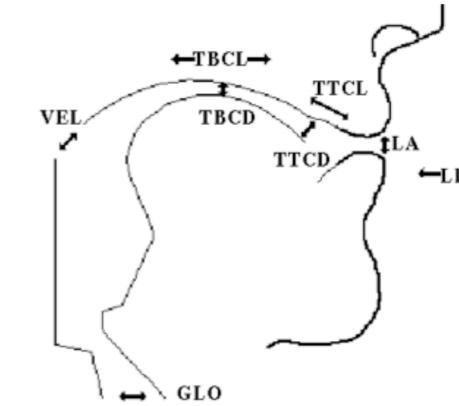
where?

how much?



Exercise: /p/

Manner of Articulation



Place of Articulation

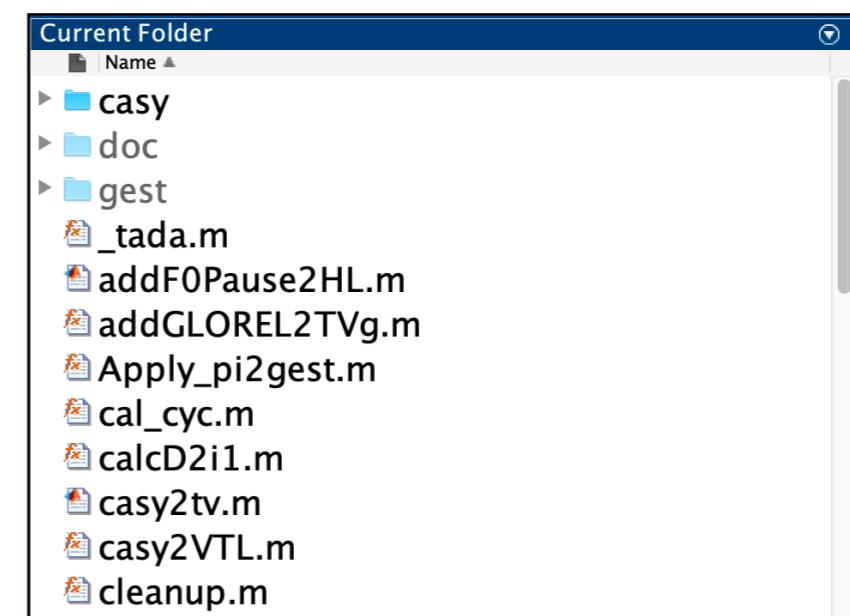
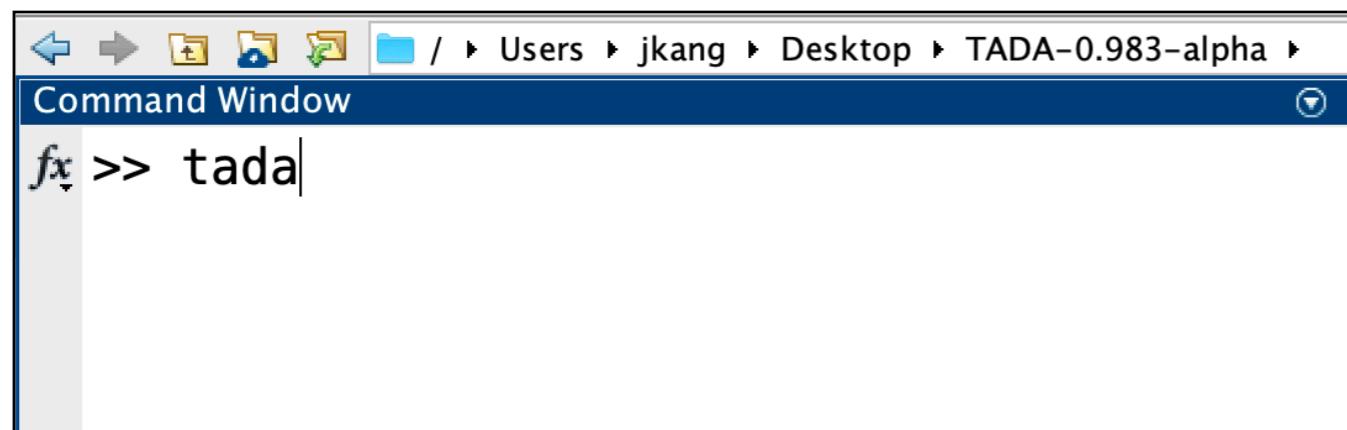
Articulators	on/off	CD	CL
Lips	on	stop	bilabial
Tongue tip	off		
Tongue body	off		
Velum	off		
Larynx	on		

Note that the active larynx ("on") means being voiceless.

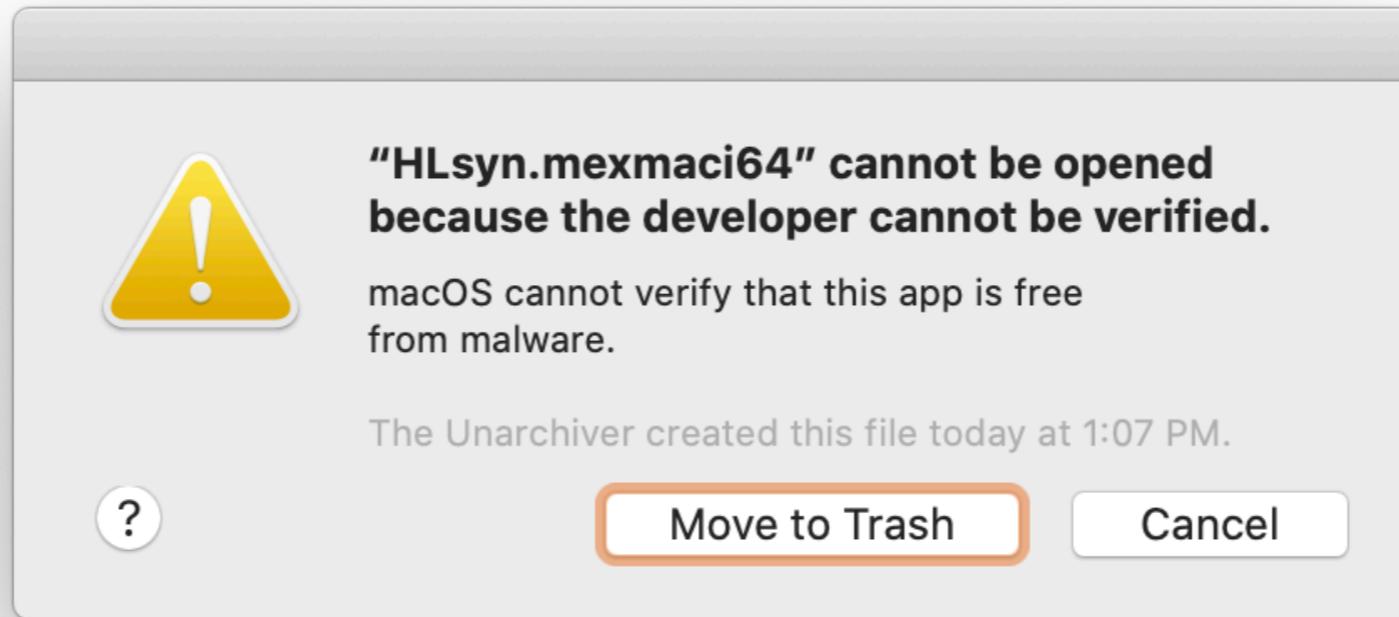
Synthesize a word “pad”

Step 0

- Download TADA.
- Unzip TADA.
- Open Matlab and go to the unzipped TADA folder.
- Type “tada” on the command line.



Warning (for mac users)

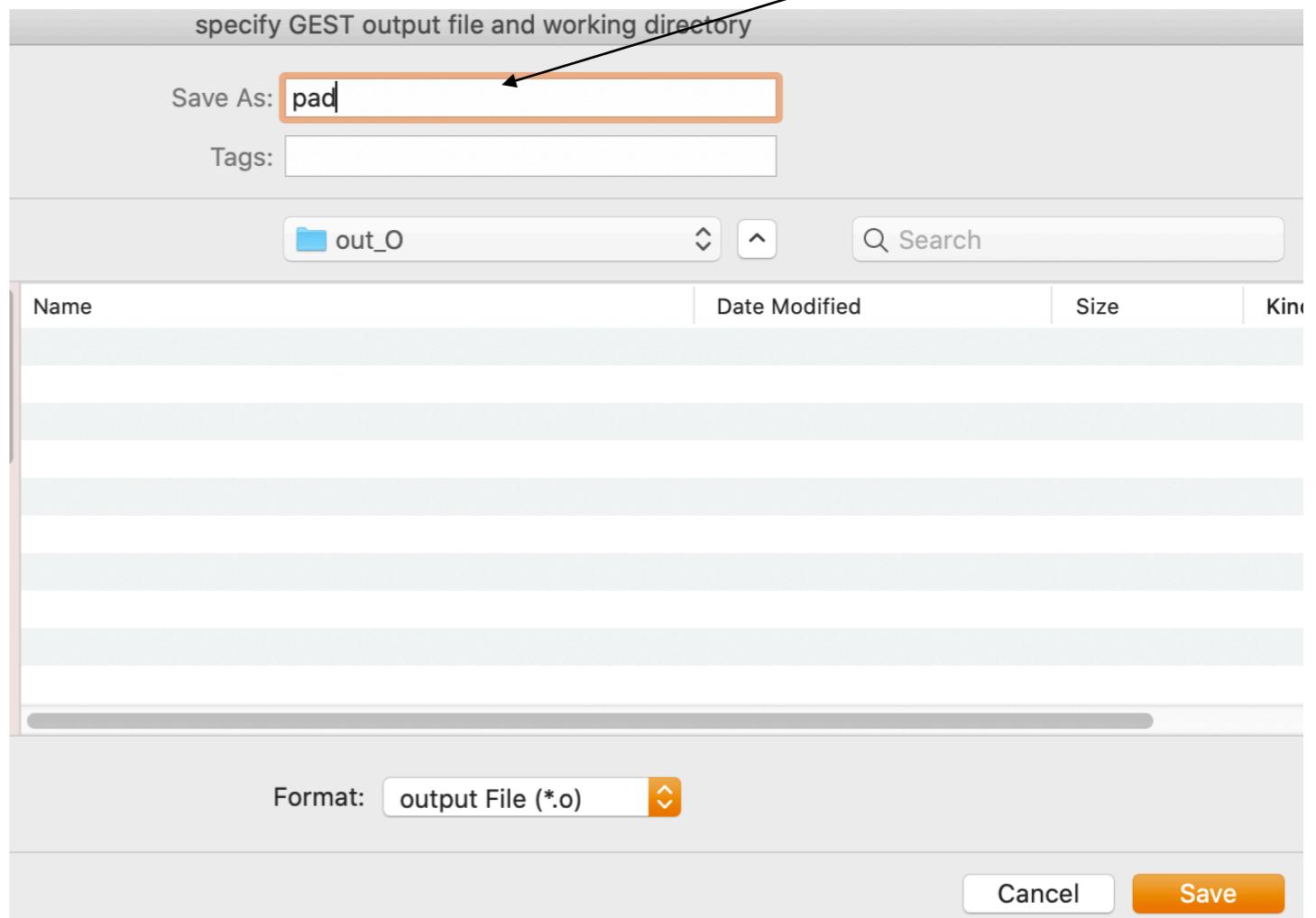
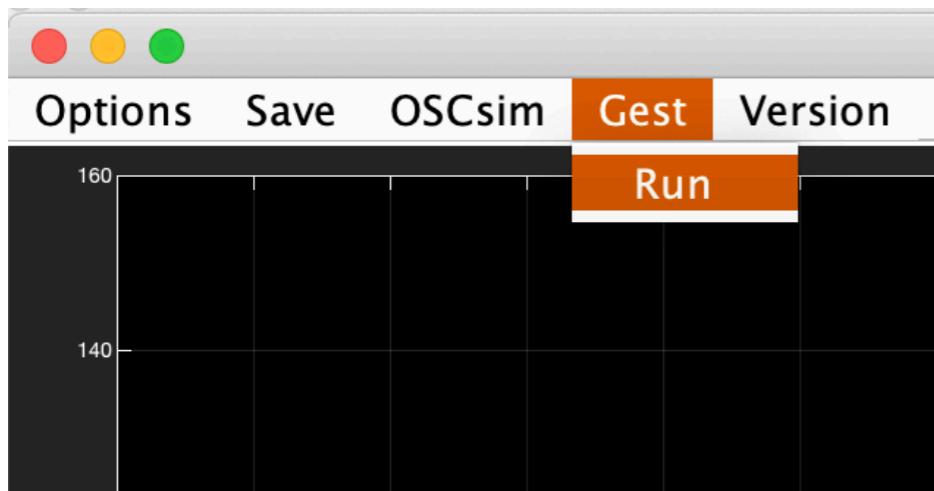


- Open Terminal.app
- Go to the TADA directory and type:
sudo xattr -r -d com.apple.quarantine ./*

(Provide your system password when asked)

Step 1

- Create *.O file by specifying a word: *Gest* → *Run*.
Coupling graph files



Gest is located on the top left corner of the TADA window.

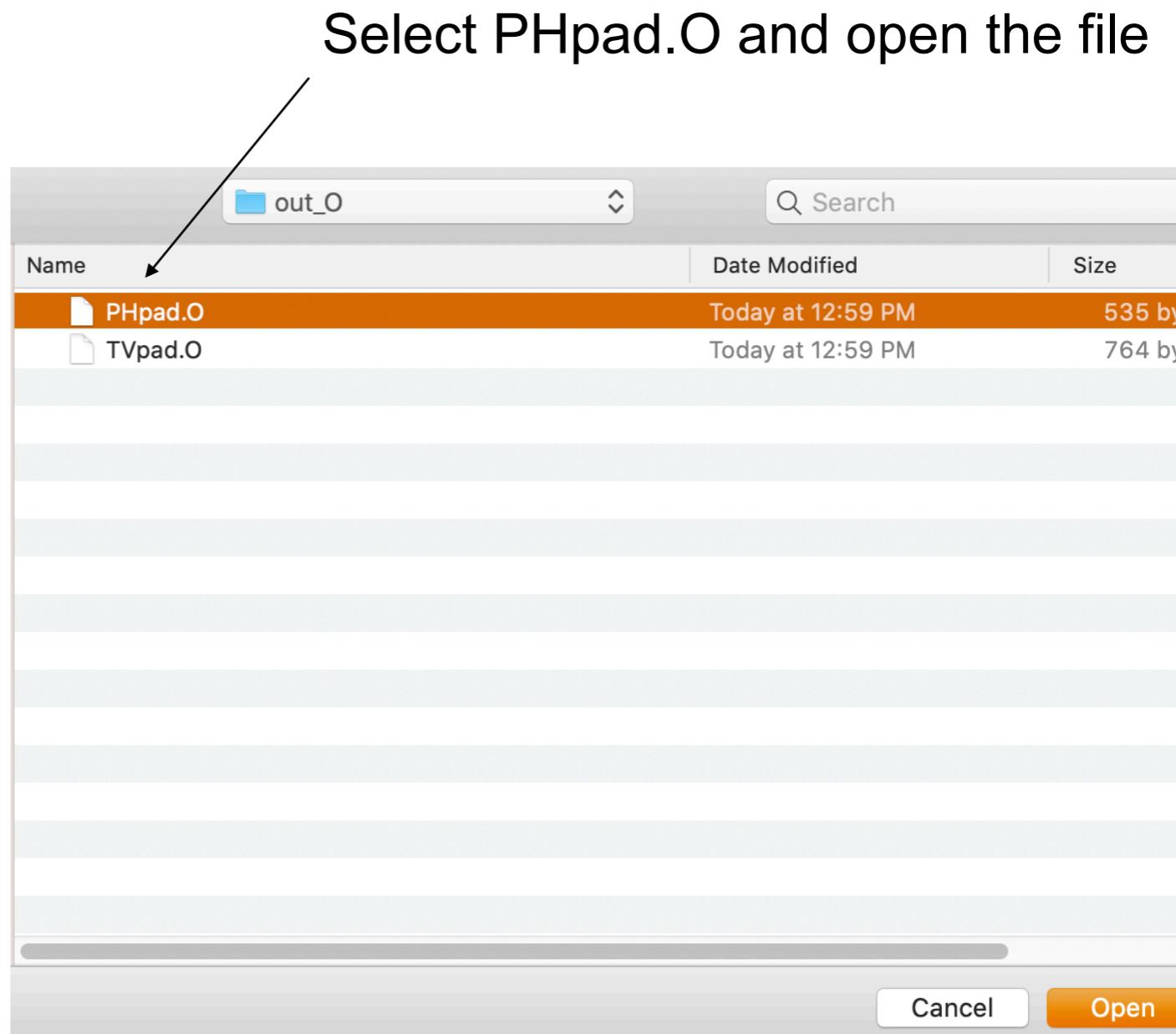
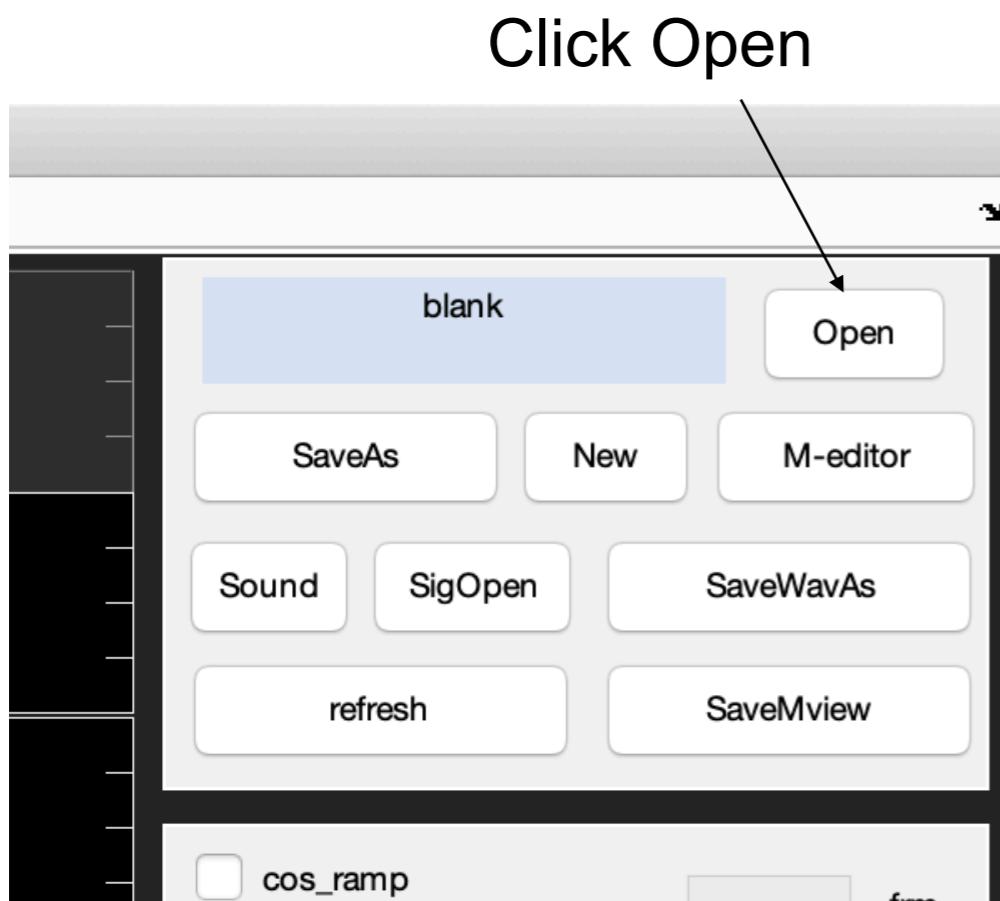
After saving, you will see another pop-up window and you can just click Okay without changing anything.

Save

12

Step 2

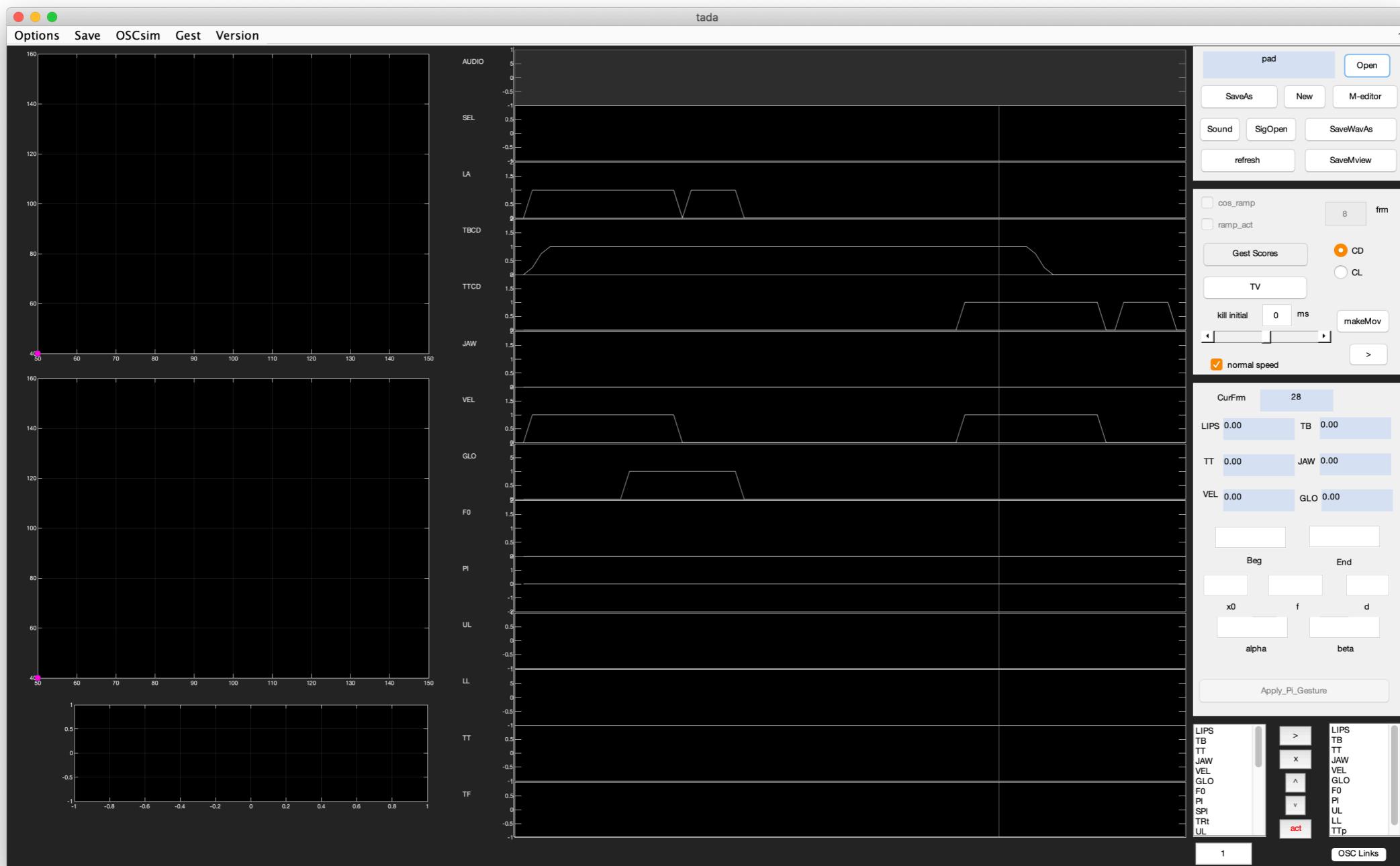
- Open *.O file.



PHpad.O file includes information about intra-gestural relationship
TVpad.O file includes information about inter-gestural relationship

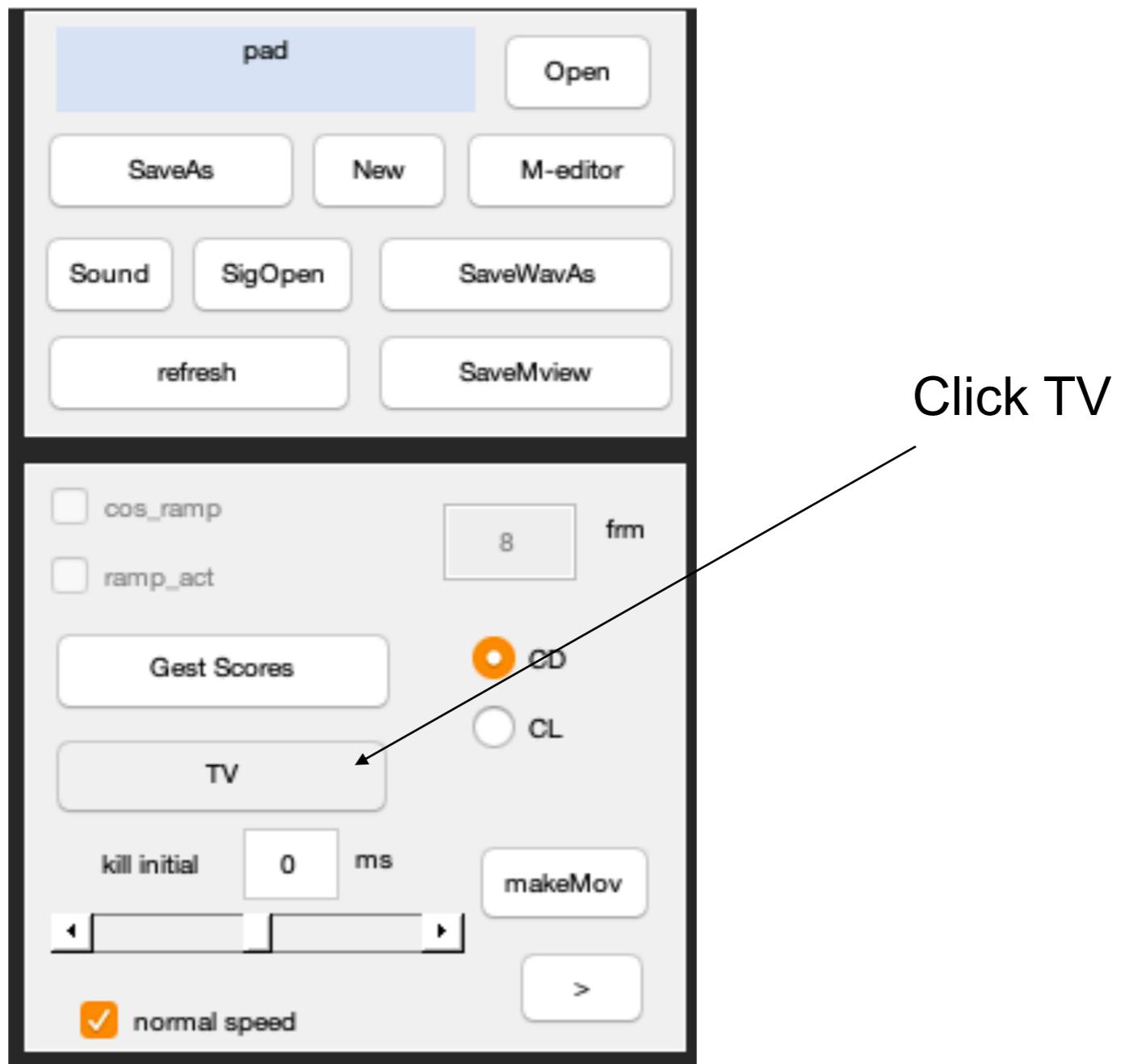
Step 2

- After opening the *.O file, gestural scores are loaded on the middle panel.



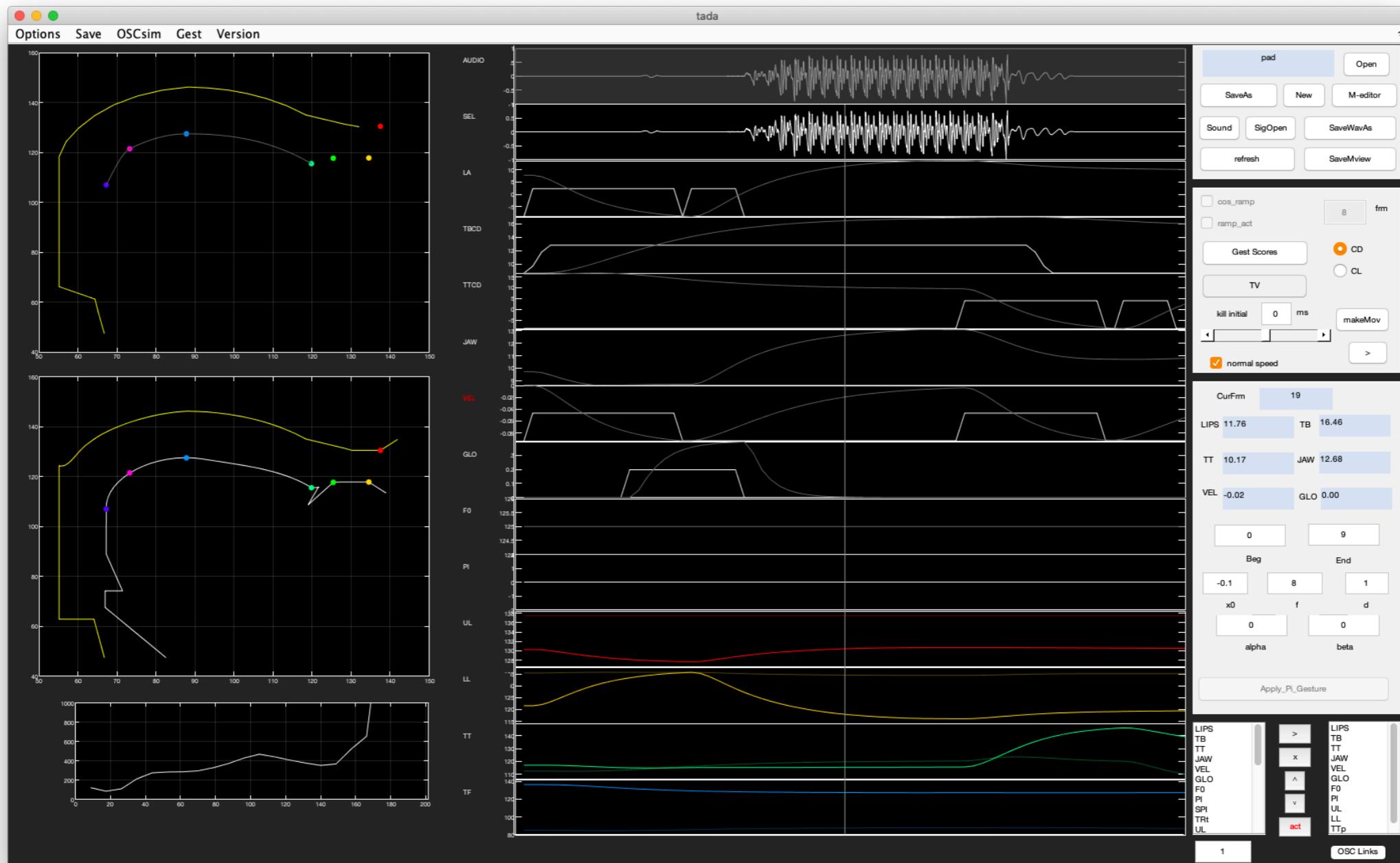
Step 3

- Press TV button.



Step 4

- Articulatory trajectories and the corresponding sound are created. Play with the cursor at the waveform and listen to the sound by clicking Sound button on the top right panel!



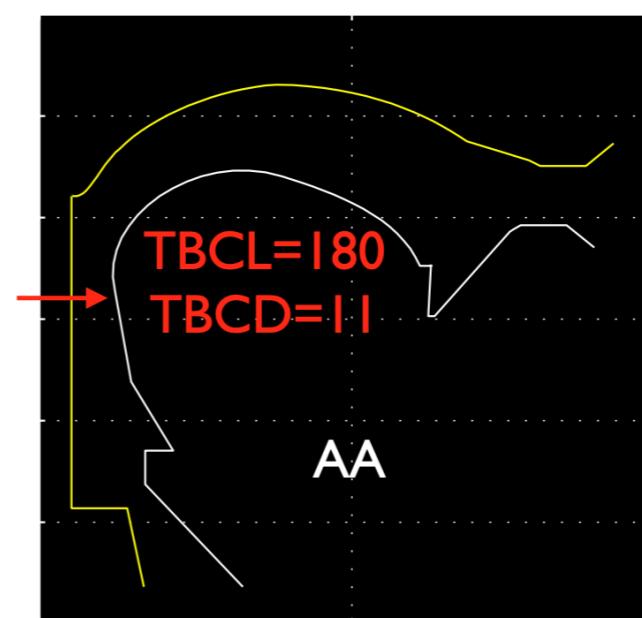
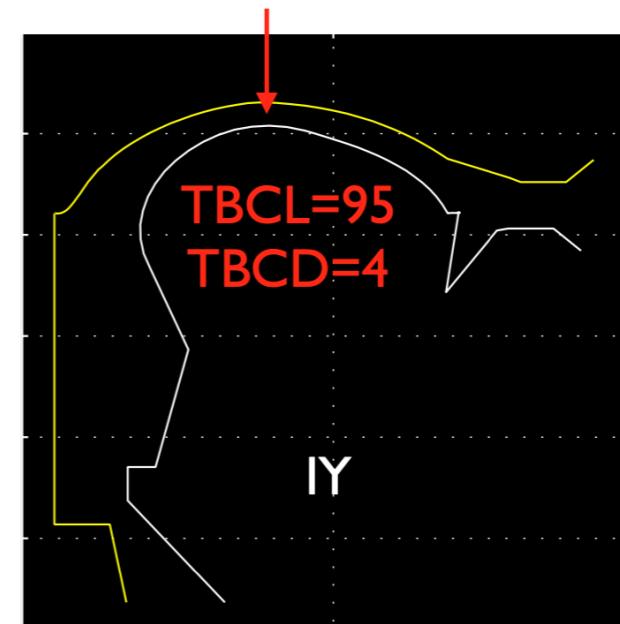
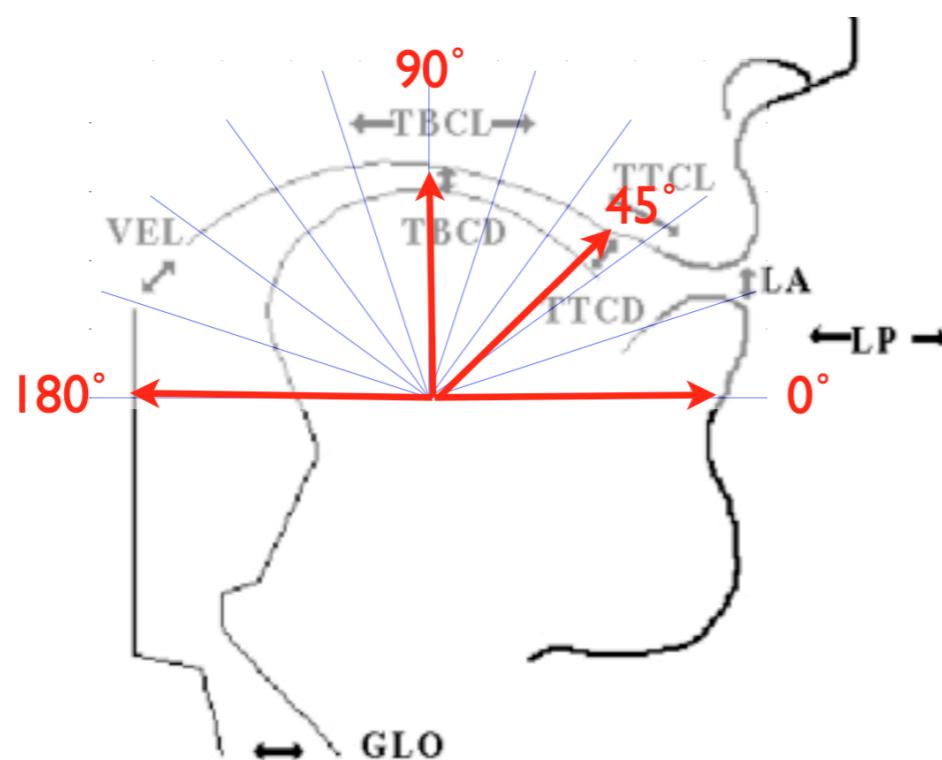
Summary

- To synthesize a word stimulus,
 - 1. Create coupling graph files (PH~.O and TV~.O)
 - 2. Create a gestural score file (TV~.G)
 - 3. Generate articulatory movement and acoustics
 - Output files will be located at
 - out_O/ coupling graph files
 - out_G/ gestural score files
 - wav/ sound output
- Already given from TADA
- ← Editable boxes
- *To clear the previous files, type:
 >> cleanup

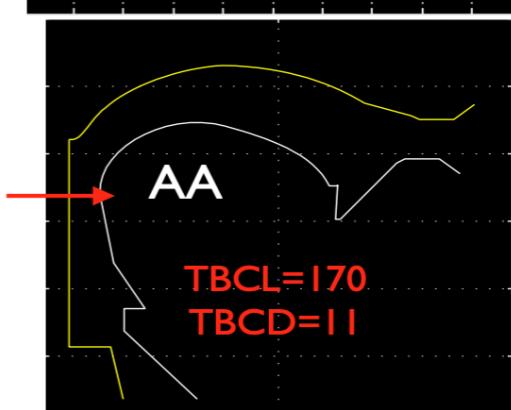
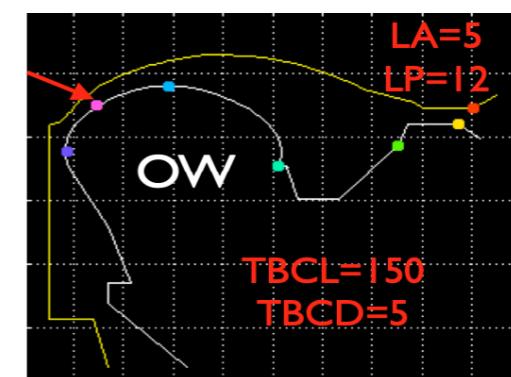
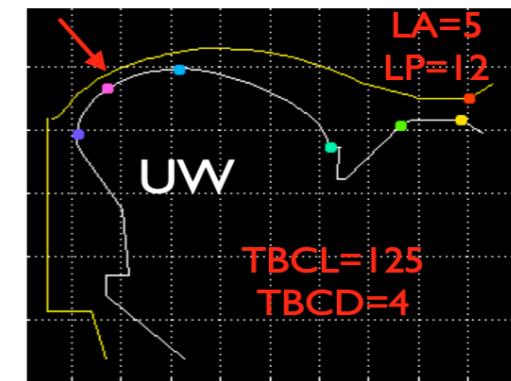
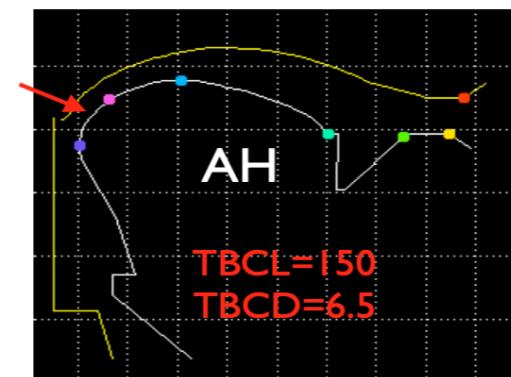
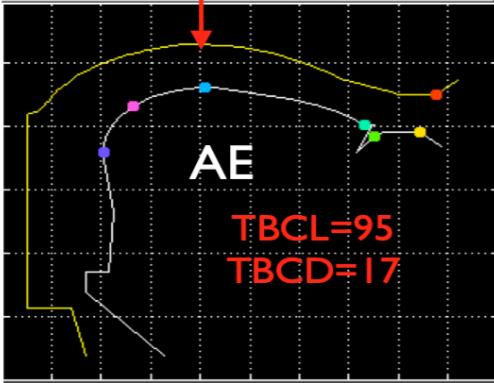
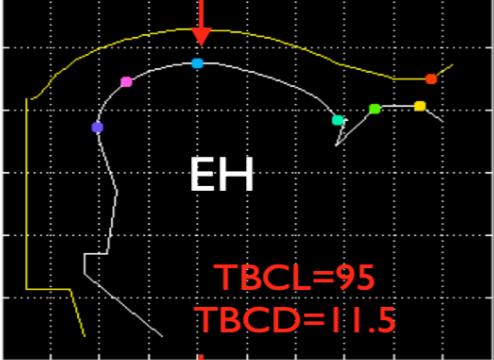
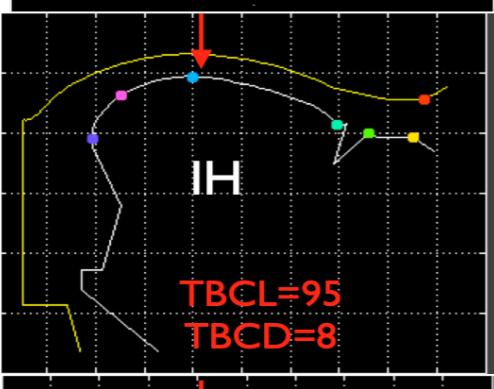
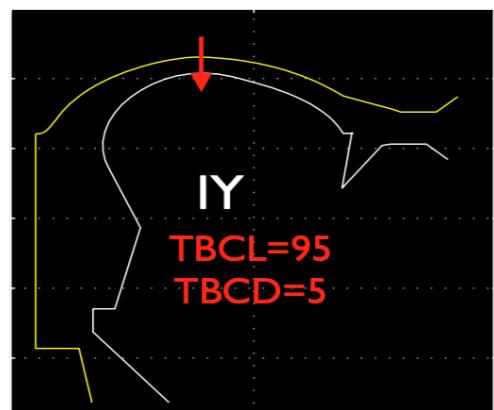
Vowel gestures

Vowel gestures

- Vowels are associated with the tongue body gesture.
 - TBCD, TBCL.

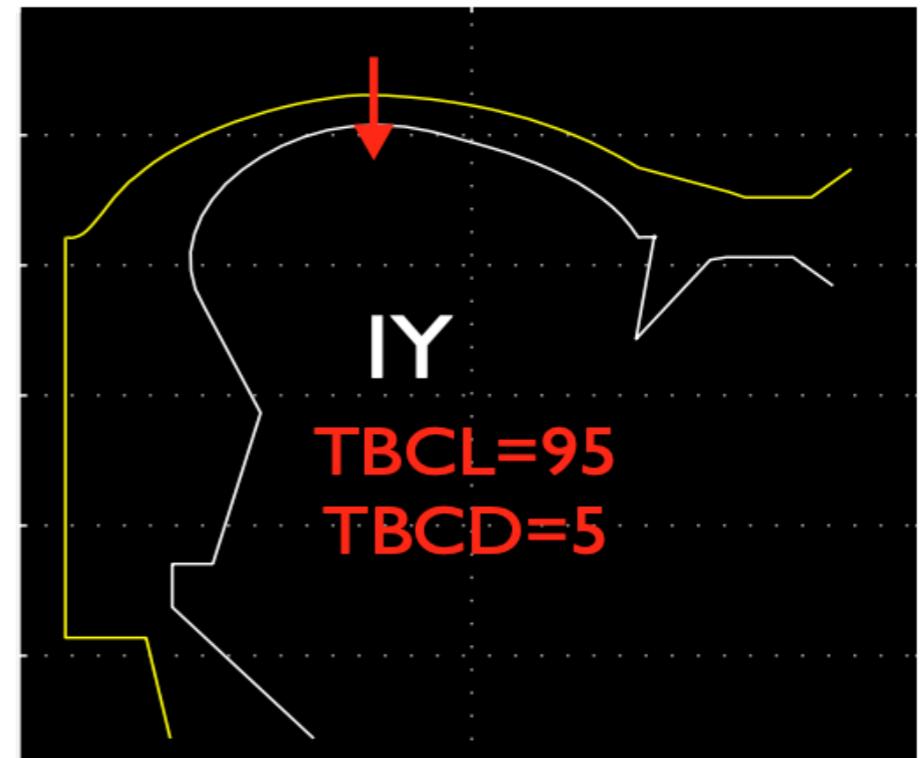
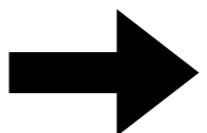
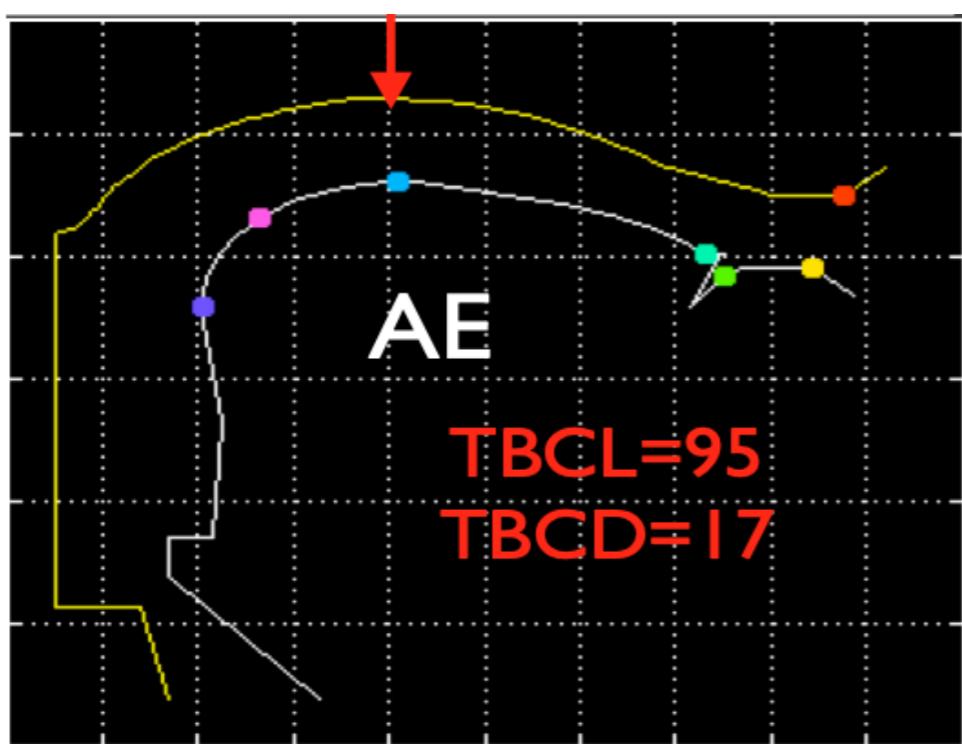


Vowel gestures



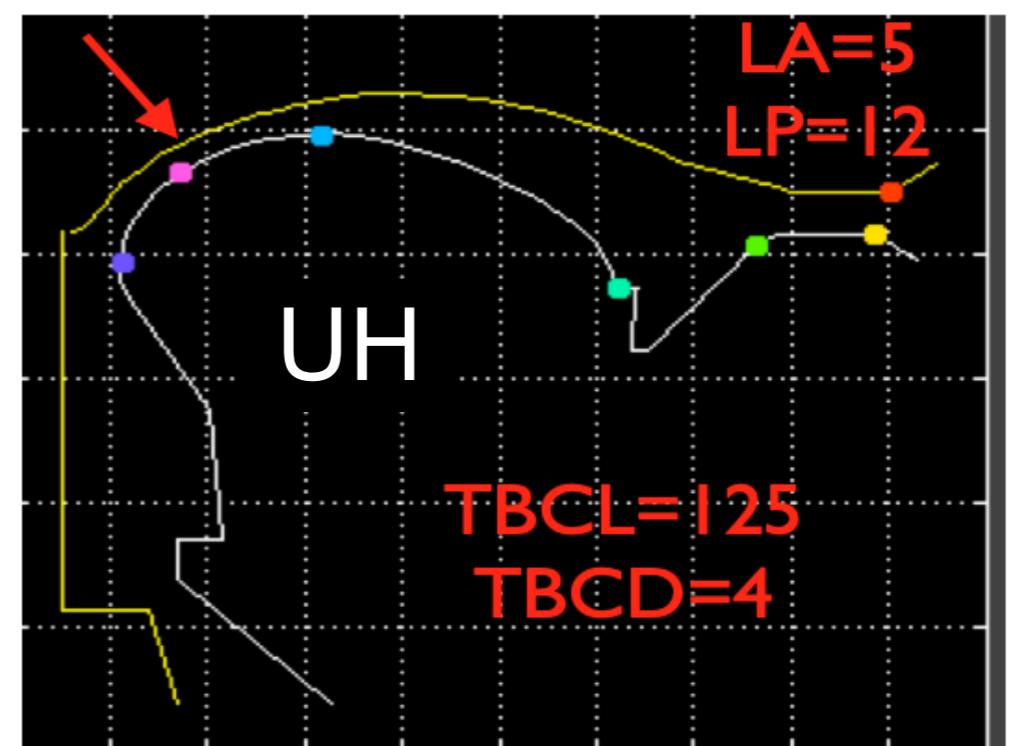
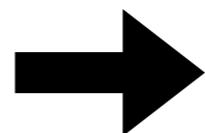
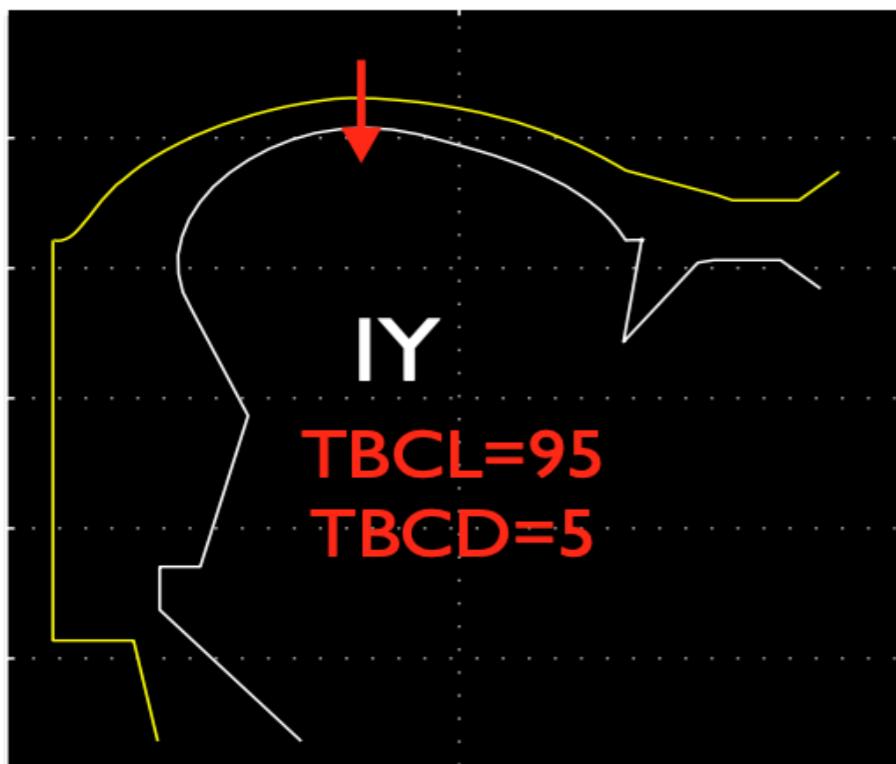
Exercise

- Synthesize “had” to “heed”.



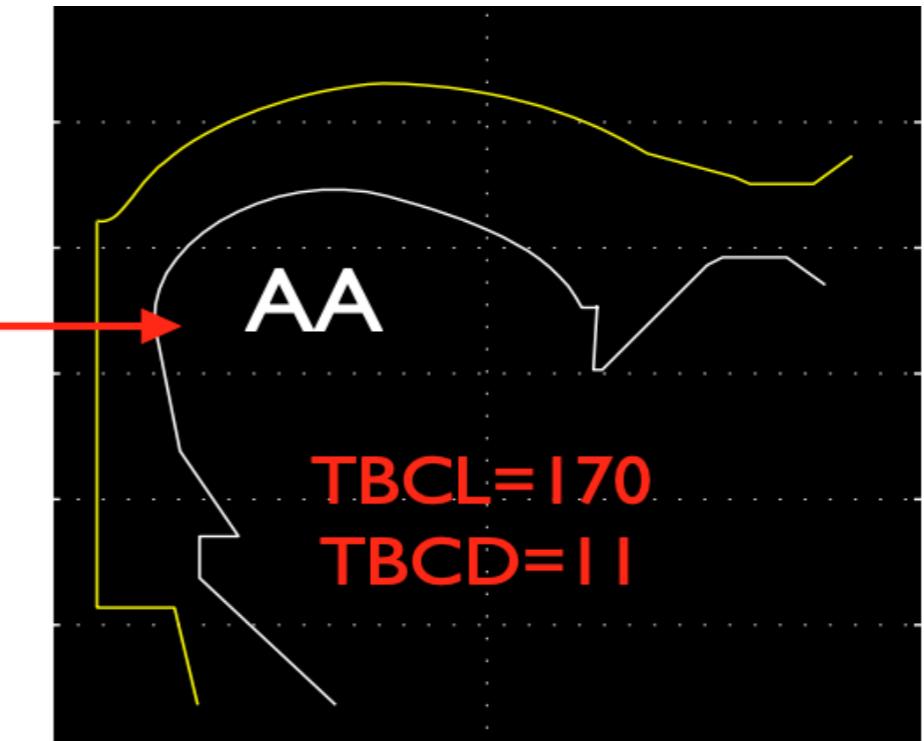
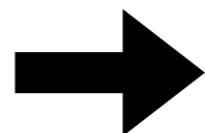
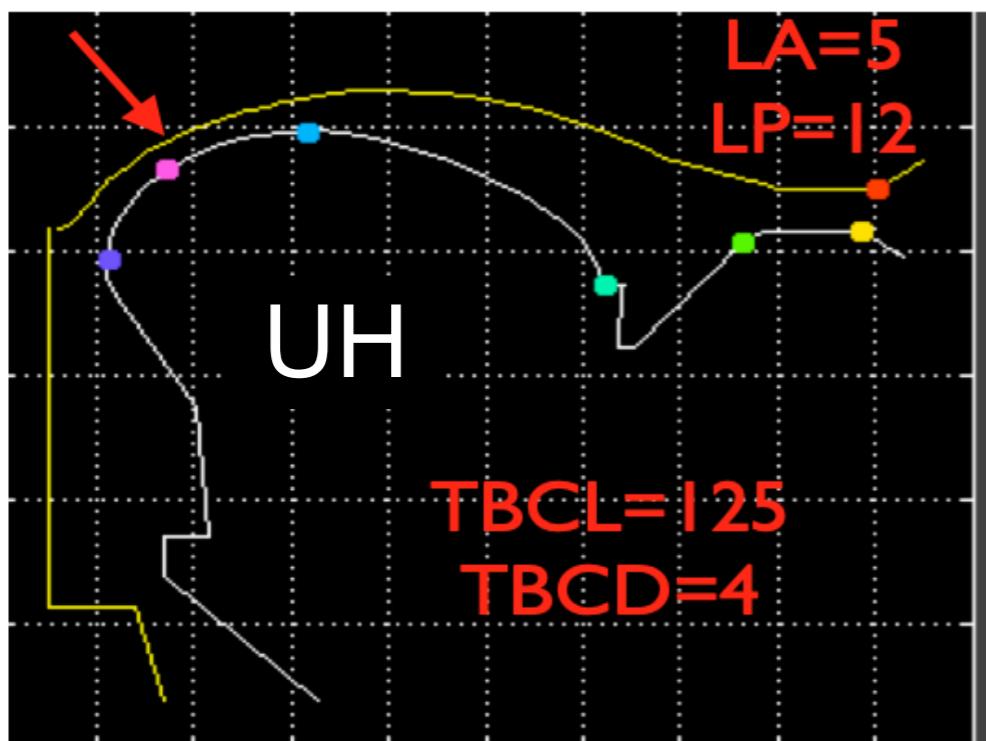
Exercise

- Synthesize “heed” to “hood”.



Exercise

- Synthesize “hood” to “hod”.



Summary

- Vowel tasks are primarily defined by TBCD and TBCL.
- LA and LP can be added for rounded vowels.

Lip Aperture

Lip Protrusion

Consonant gestures

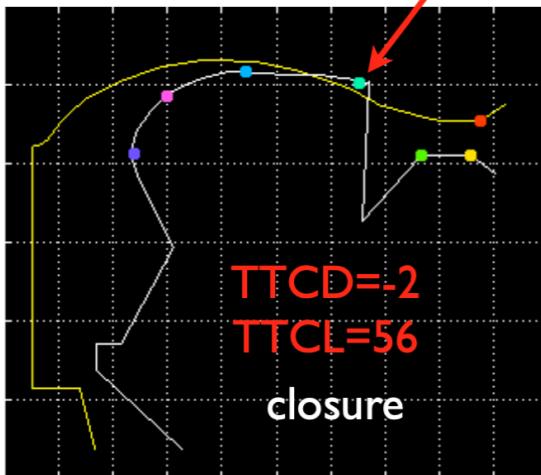
Consonant gestures

- Making consonants involve more articulators and tasks than vowels.
 - TB, TT, VEL, LIP, JAW (with GLO)
-

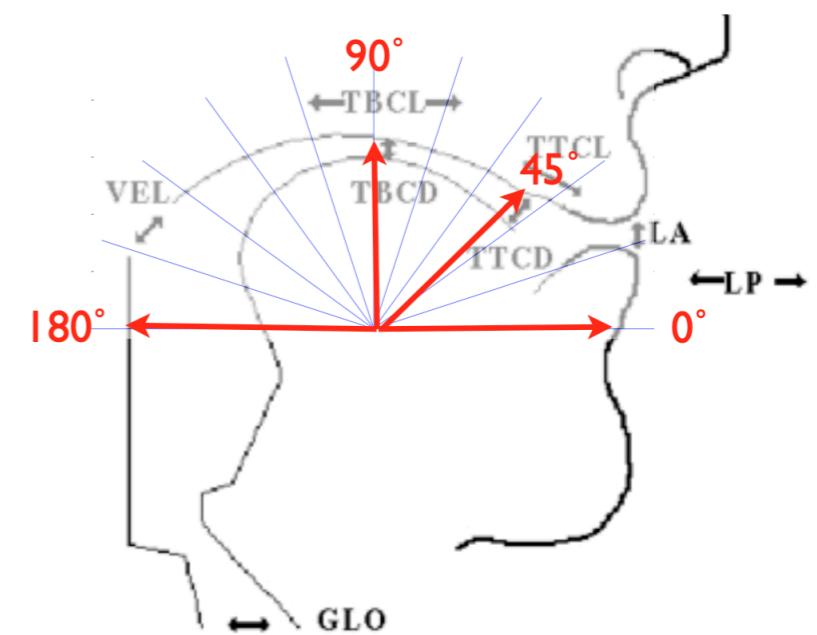
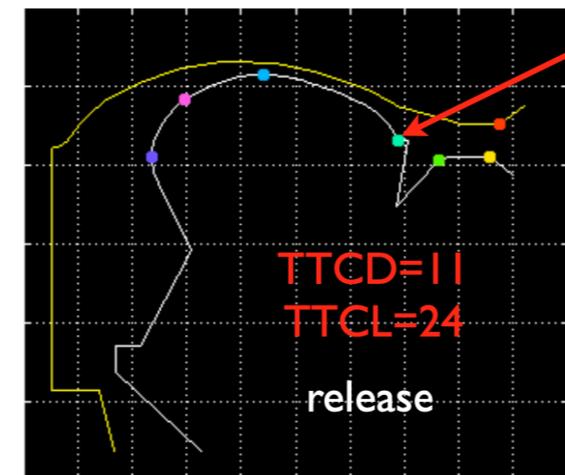
Tongue Tip Tasks

for /t, d, n/

Closing gesture

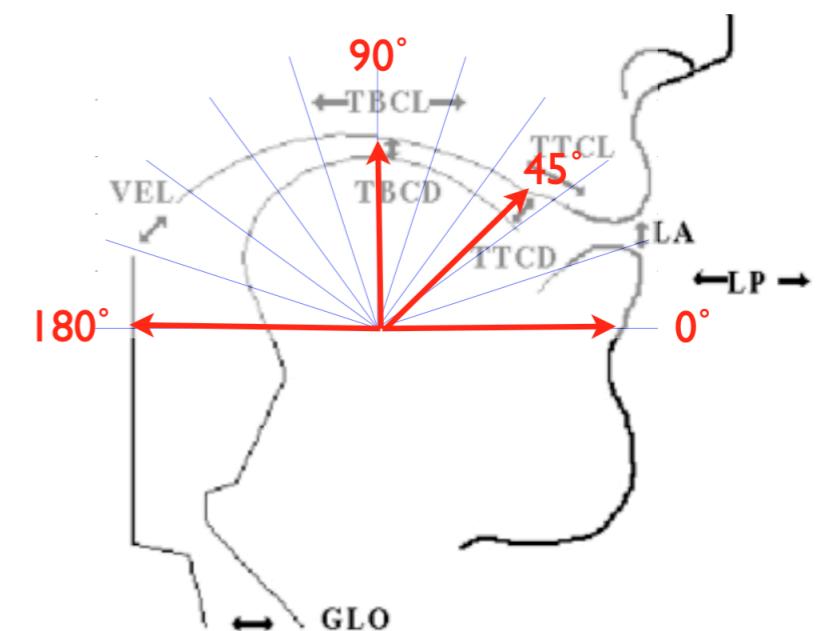
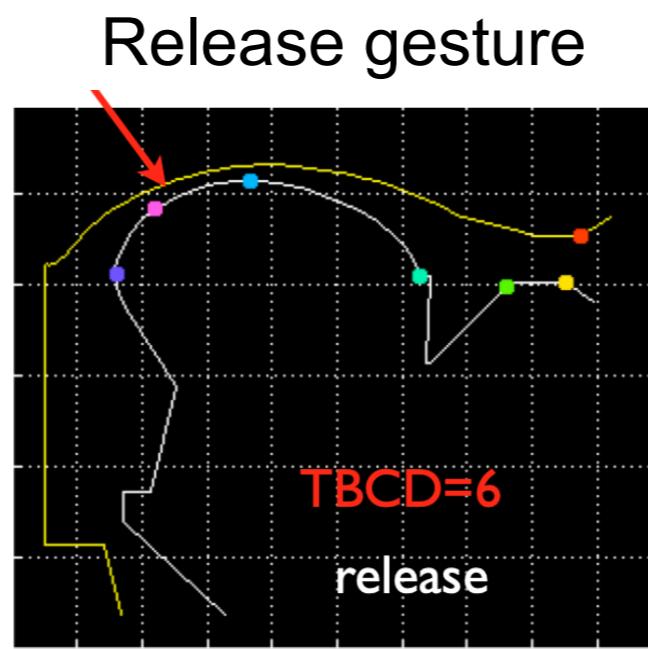
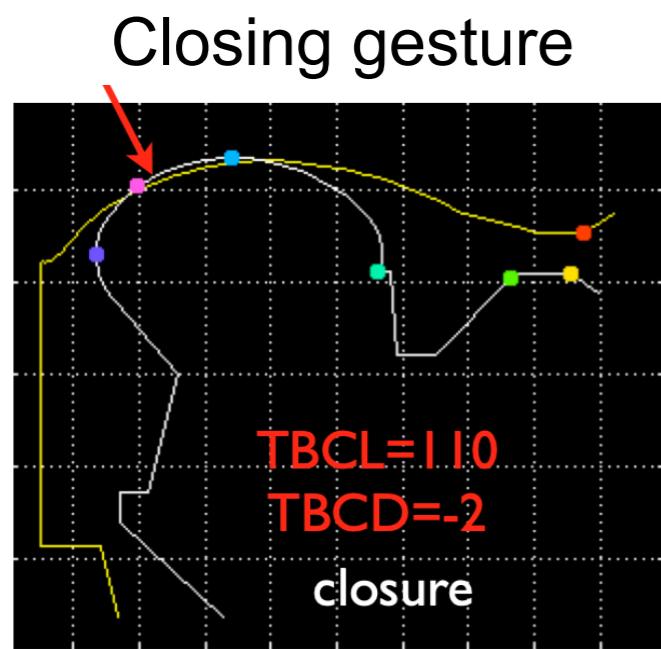


Release gesture



Consonant gestures

Tongue Body Tasks
for /k, g, ɳ/

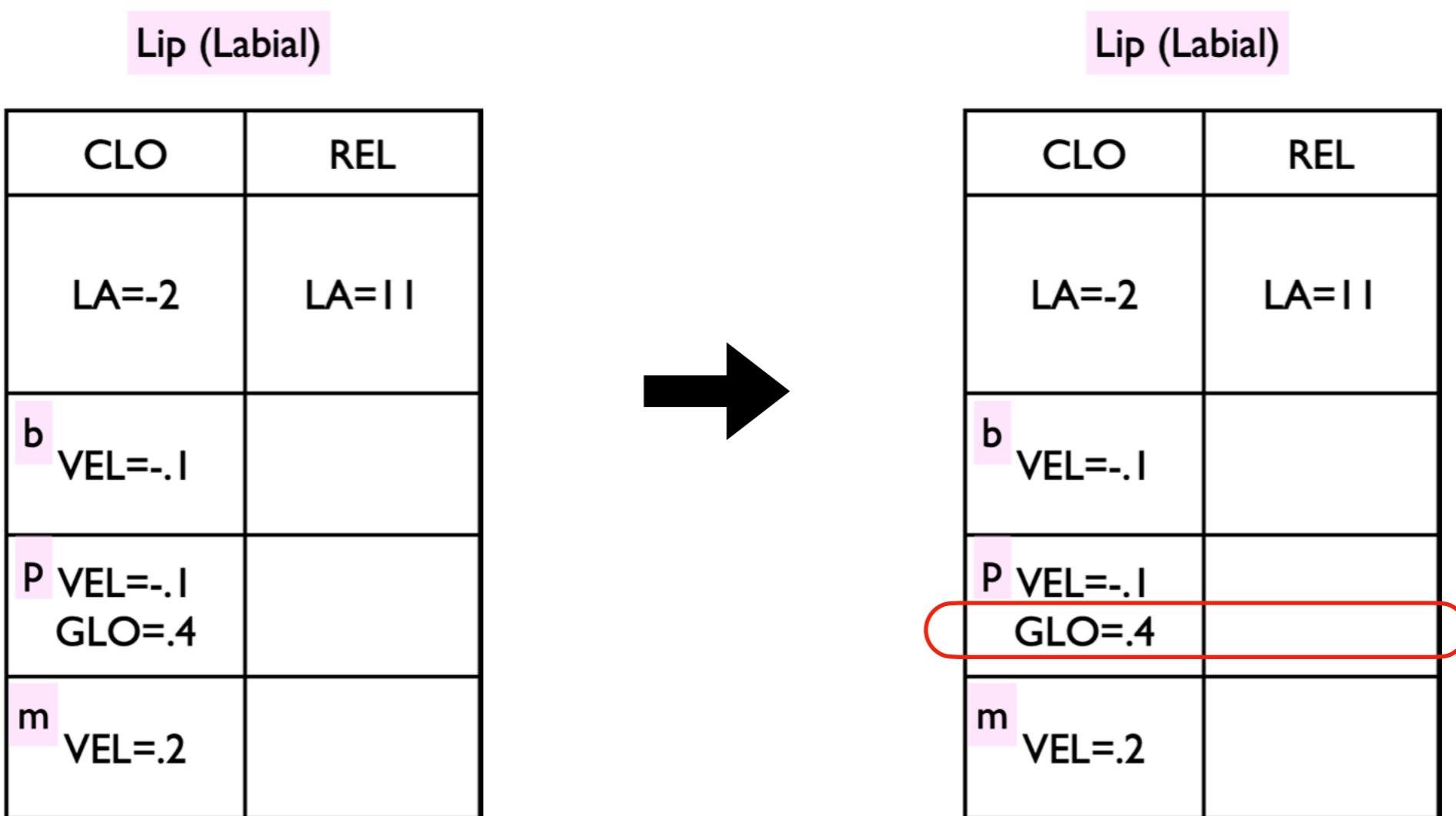


Tasks for stop consonants

Lip (Labial)	T Tip (Alveolar)	T Body (Velar)			
CLO	REL	CLO	REL	CLO	REL
LA=-2	LA=11	TTCL=56 TTCD=-2	TTCL=24 TTCD=11	TBCL=110 TBCD=-2	TBCD=6
b VEL=-.1		d VEL=-.1		g VEL=-.1	
p VEL=-.1 GLO=.4		t VEL=-.1 GLO=.4		k VEL=-.1 GLO=.4	
m VEL=.2		n VEL=.2		ŋ VEL=.2	

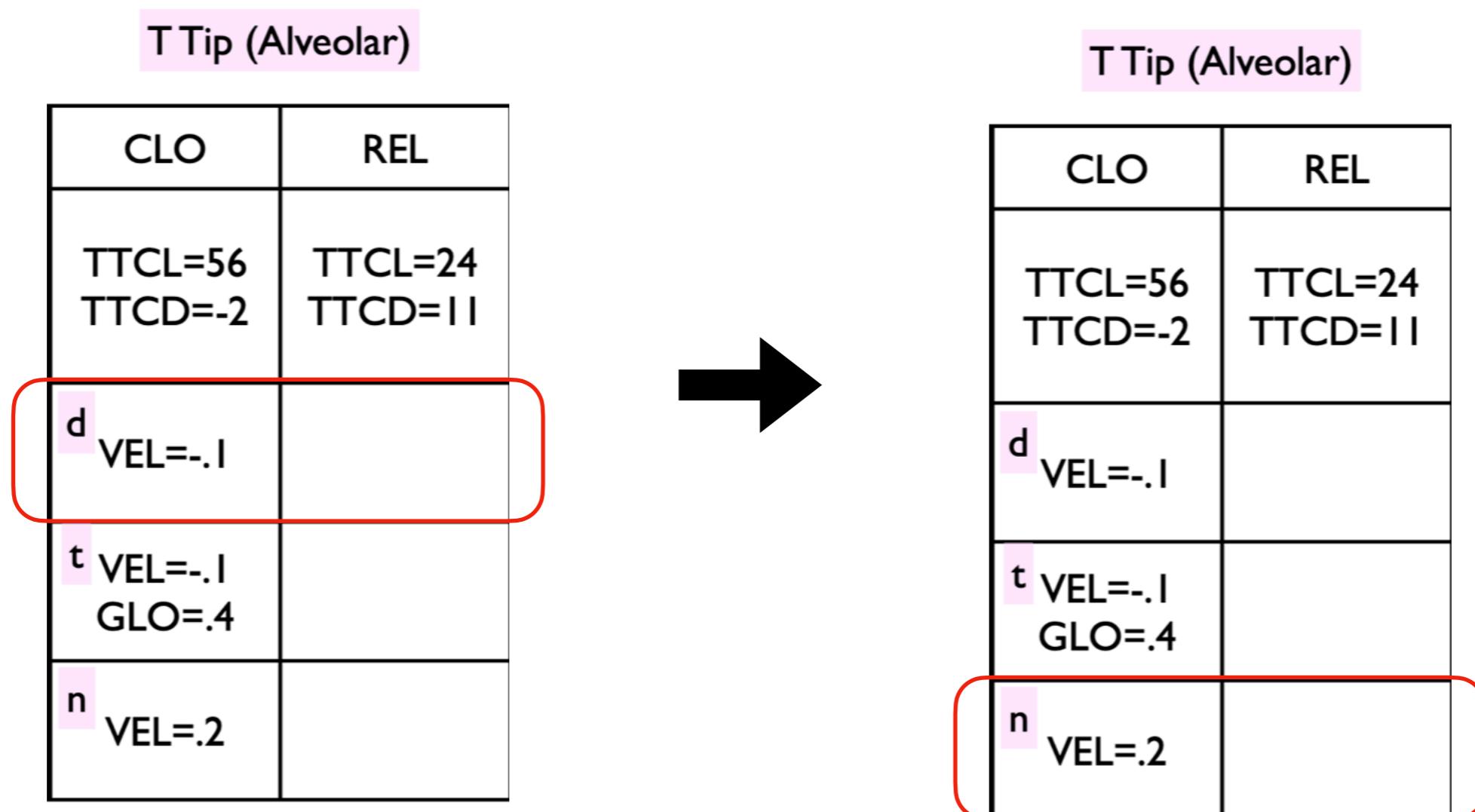
Exercise

- Synthesize “pit” to “bit”.



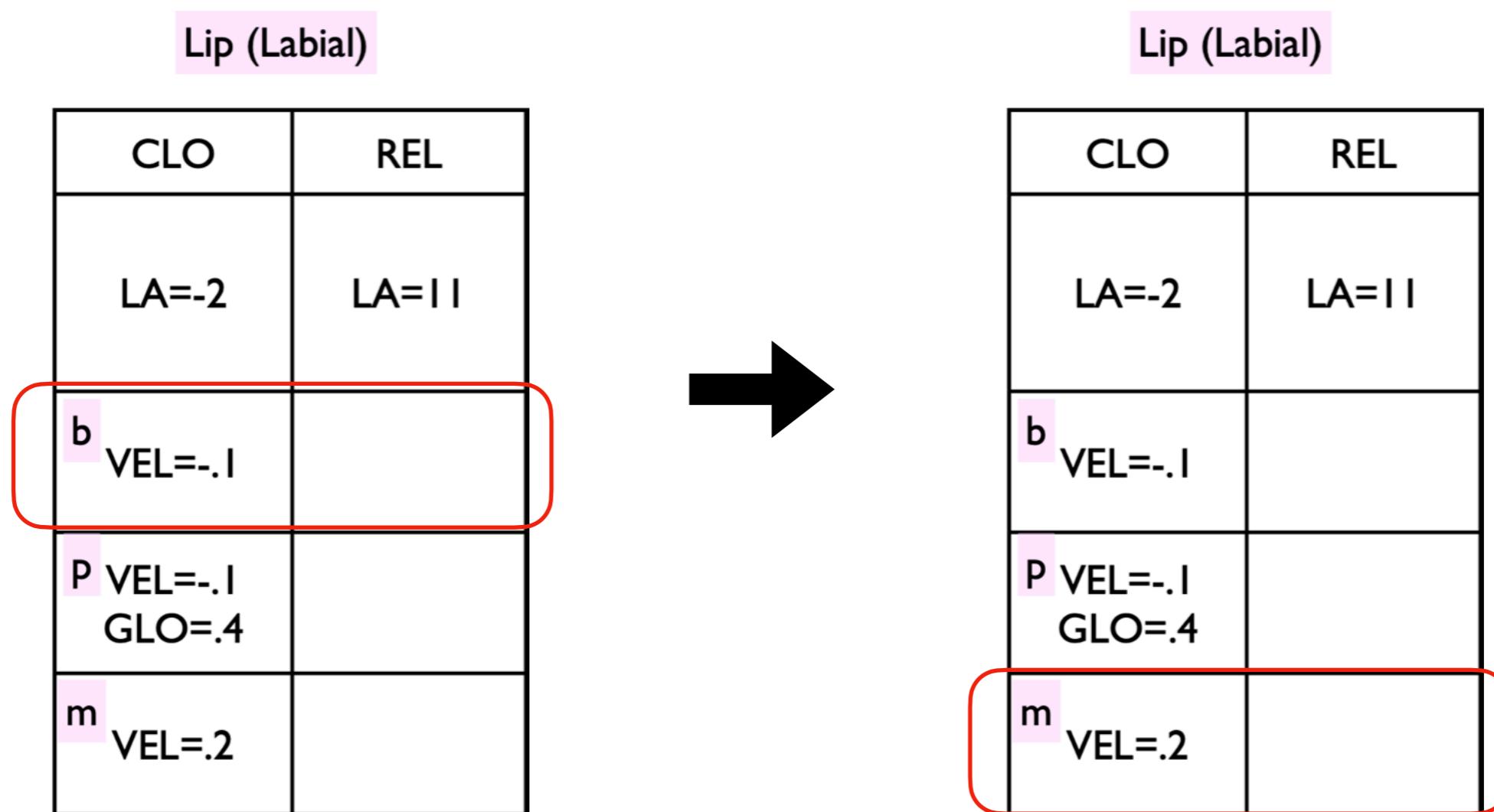
Exercise

- Synthesize “bad” to “ban”.



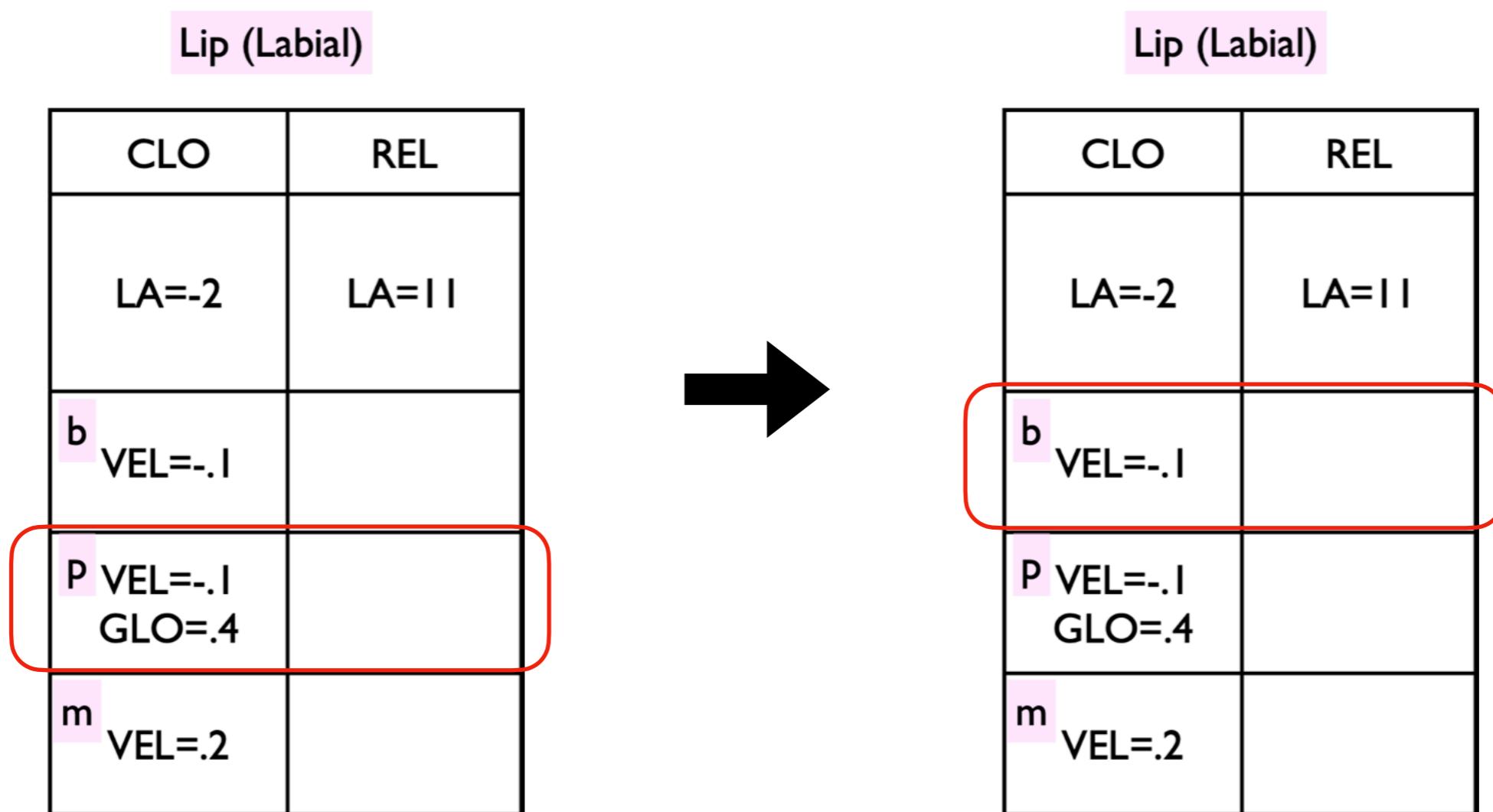
Exercise

- Synthesize “ban” to “man”.



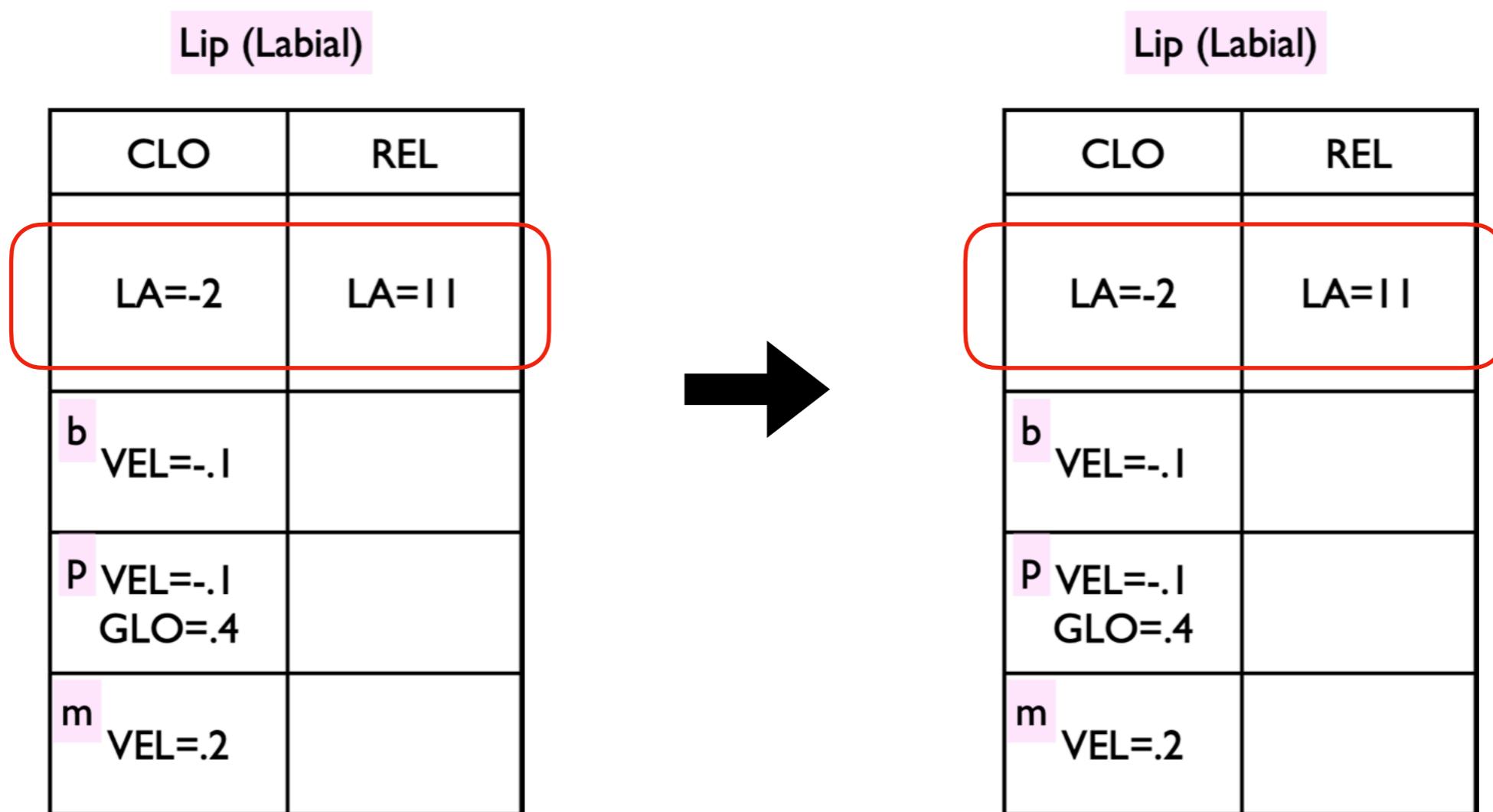
Exercise

- Synthesize “pad” to “bad”.



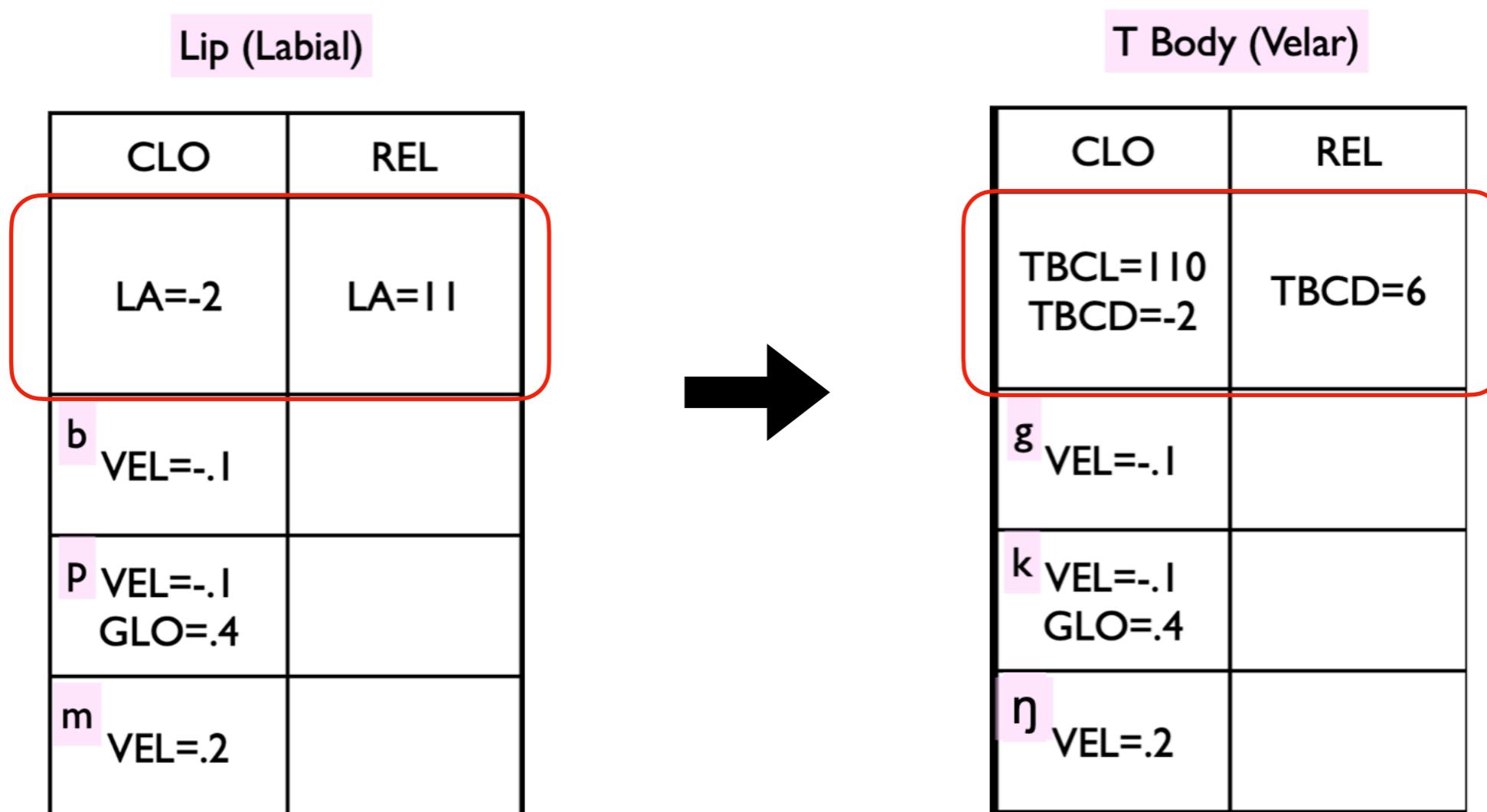
Exercise

- Synthesize “bad” to “wad”.



Exercise

- Synthesize “pit” to “kit”.

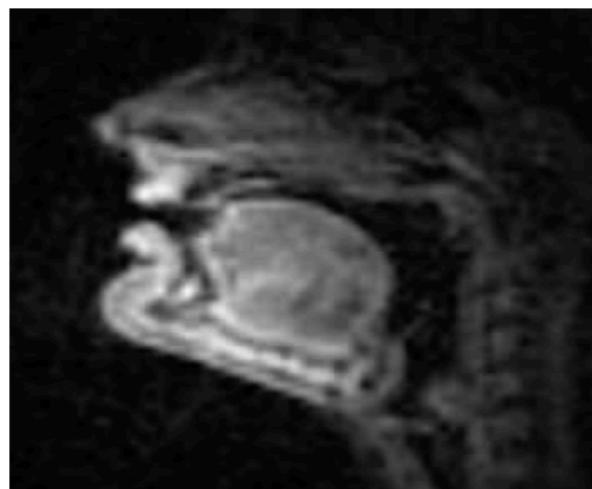


Summary

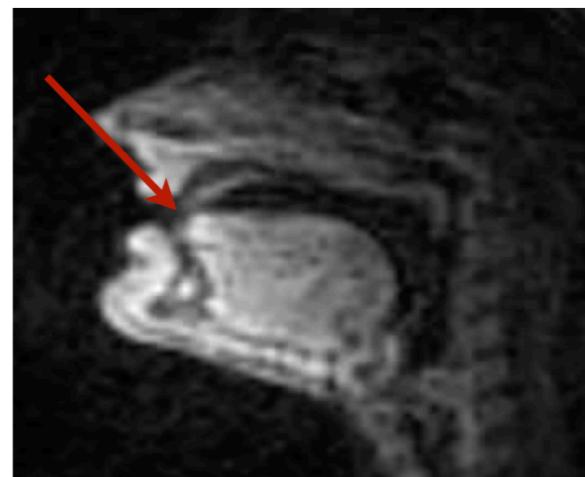
- Consonant tasks can be manipulated by changing constrictions at the lip, tongue tip and tongue body.
- Changing the release gesture for a stop can change the VOT interval.

Fricatives

- Fricatives are implemented as oral constriction tasks.



θ



z



ʃ

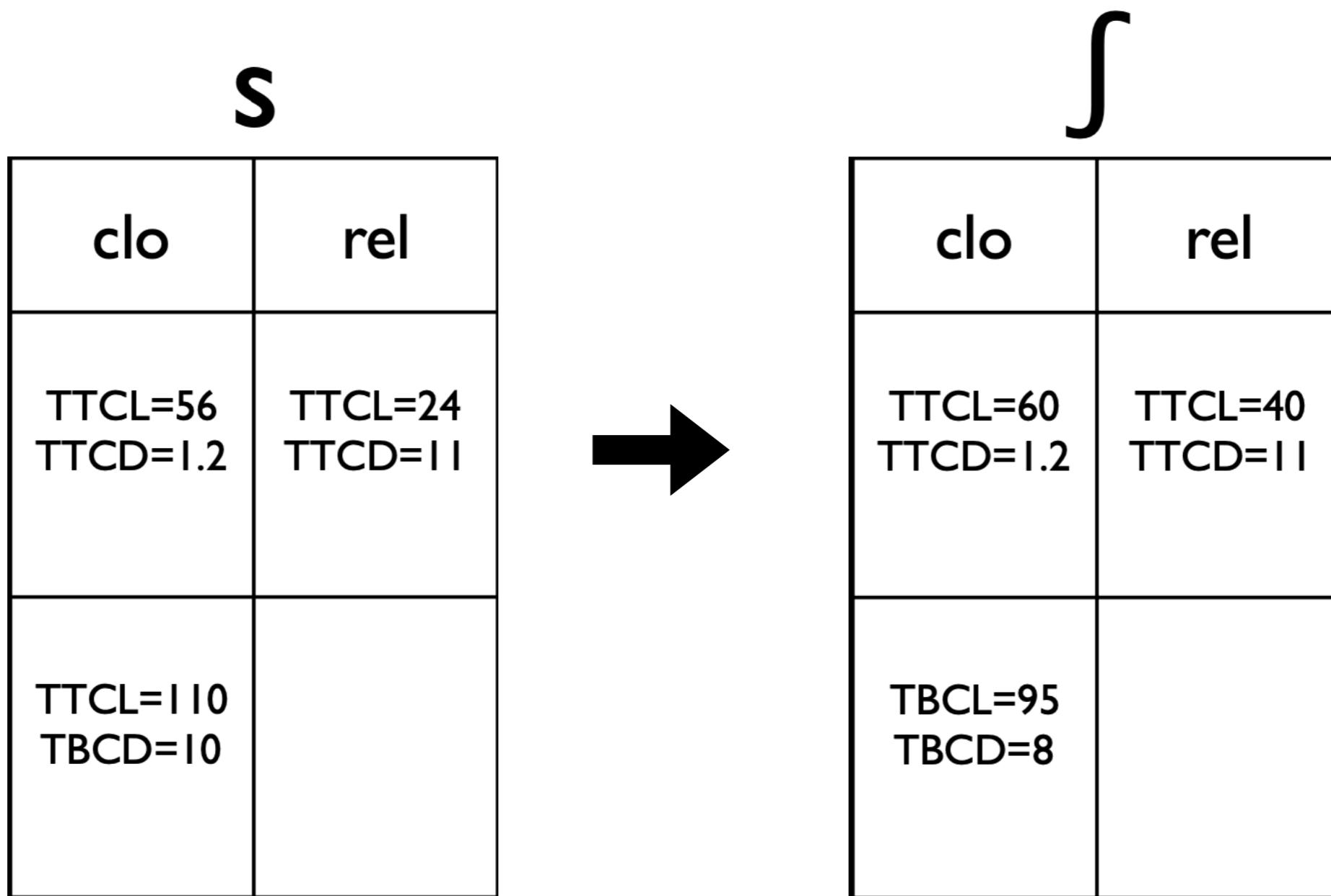


Fricatives

		θ	S		\int		
		clo	rel	clo	rel	clo	rel
TT		TTCL=40 TTCD=1.2	TTCL=24 TTCD=11	TTCL=56 TTCD=1.2	TTCL=24 TTCD=11	TTCL=60 TTCD=1.2	TTCL=40 TTCD=11
	TB			TTCL=110 TBCD=10		TBCL=95 TBCD=8	

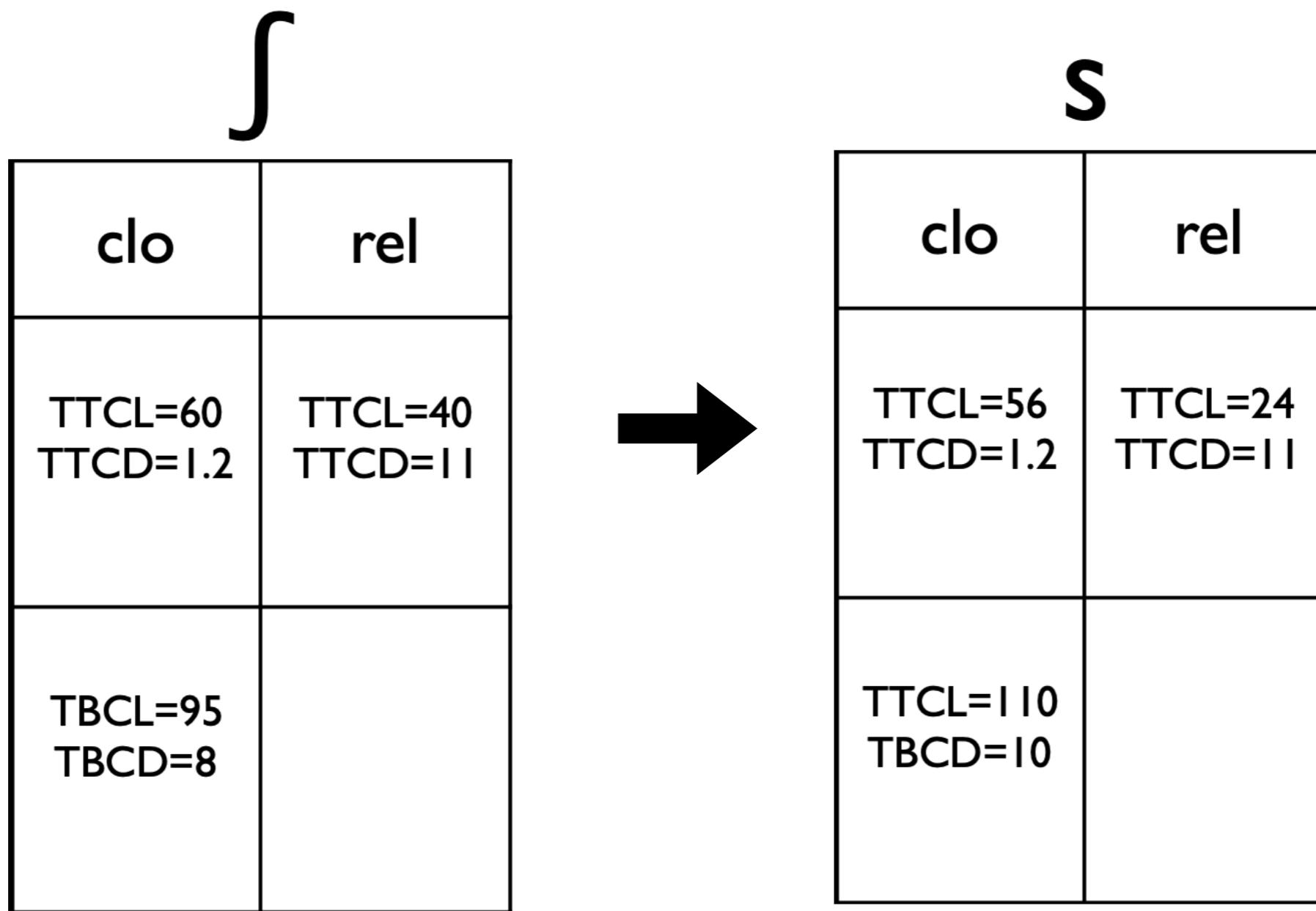
Exercise

- Synthesize “see” to “she”



Exercise

- Synthesize “shell” to “sell”



Gestural overlap and timing

- bad ban vs bad#ban

Gestural overlap and timing

- Slower/faster consonant constrictions in “bad ban”
- “ban” to “ba:n”

Conclusion

- TADA is a useful tool for understanding gestural organization of speech by simulating dynamics of gestures.
 - Vowel and consonant gestures (TT, TB, UL, LL, VEL, GLO)
 - Dynamical parameters (target, stiffness, damping)
 - Gestural timing and overlap

Limitations

- Peri-sagittal specification of vocal tract.
- Aerodynamic modeling of speech.
- Naturalness of the sound output.

Final remarks

- Try one your own!
 - Exercises
 - Matlab