TASK DYNAMIC APPLICATION

Dzhuma Abakarova

Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), University of Potsdam

dzhuma.abakarova@uni-potsdam.de

Summer school 2022: Coping with the complexity in speech production and perception

Chorin, July 8

GOALS

- You know what TADA is
- You know what TADA can be used for
- You are able to use TADA for simple manipulations
- You know where to learn more about TADA

OUTLINE

1. Introduction
    1. Why model speech production
    2. Why TADA
    3. What can you do with TaDA (examples)
2. How to do something in TaDA
    1. GUI demonstration
    2. Scripts
3. Tiny simulation experiment with TaDA
    1. (and how to report it)

Introduction

TASK DYNAMICS APPLICATION (TADA)

- a computational model of speech production
- an implementation of Articulatory Phonology and Task Dynamic (AP/TD) approaches to speech production

Developed by (alphabetical)

Cathe Browman
Louis Goldstein
Hosung Nam
Michael Proctor
Philip Rubin
Elliot Saltzman
Mark Tiede

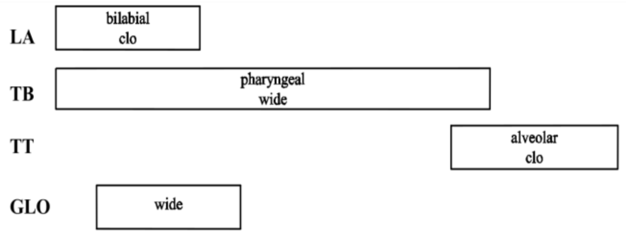
TASK DYNAMICS APPLICATION (TADA)

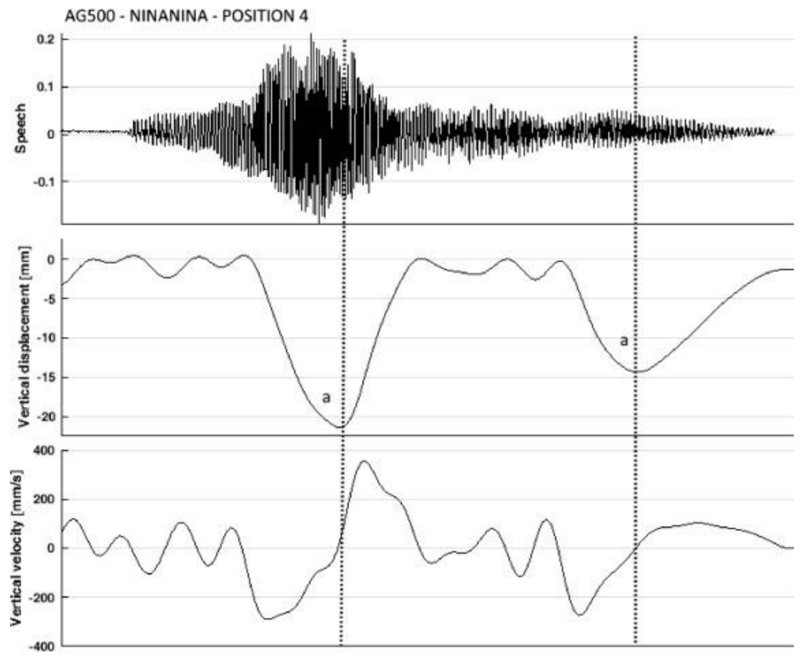
- Download
- TADA demo video singing "The Star-Spangled Banner“
- Manual

TADA

„model“

TADA





Why model speech production

when it is either impossible or impractical to create experimental conditions in which scientists can directly measure outcomes.

Examples for speech research:

- Difficult to record populations
- Complex interactions
- Long-term effects

Why model speech production

Experiment

Reduces complexity of the environment

Model

Reduces complexity of the environment

AND

The organism (the system)

Why model speech production

Models serve to

1. infer the mechanisms underlying observed responses by reducing complexity
2. encapsulate current theoretical understanding
3. organize and integrate empirical findings
4. force us to specify our assumptions
5. generate predictions

Which model

Models of speech production – model at least

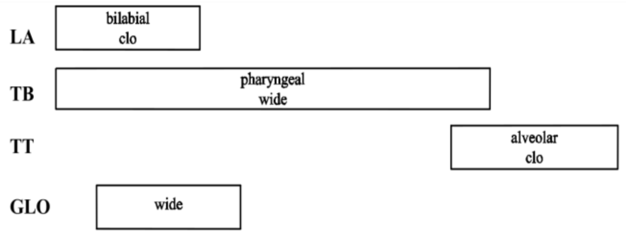
- DIVA (Guenther, 1994)
- TADA (Saltzman and Kelso, 1987; Saltzman and Munhall, 1989)
- SFC (Houde and Nagarajan, 2011)
- FACTS (Parrell *et al.*, 2018)
- ACT (Kröger *et al.*, 2009)
- GEPPETO (Perrier *et al.*, 2006)

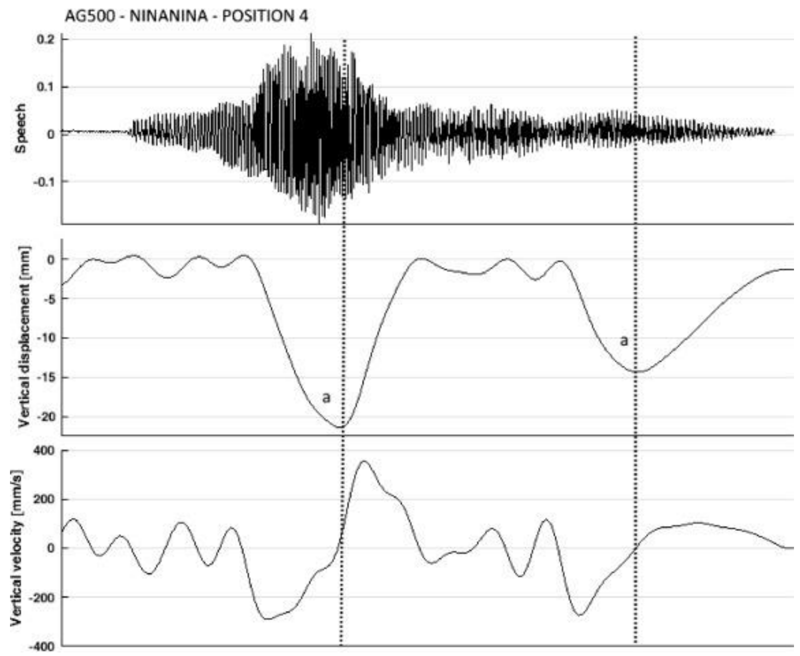
Why TADA

- Dynamic vocal tract articulator movements
- Models CV coarticulation well
- Output easy to compare to articulatory data: EMA, MRI, ultrasound
- Meant for testing hypotheses
- Requires little expertise in programming for running simulation experiments

Why TADA

„model“ → TADA





Why TADA

TADA: possible uses

- Generate perceptual stimuli with known (pseudo-)articulatory properties.
- Test hypothesized causes for regularities in speech production kinematics.
    - syllable structure effects
        - relative timing of articulatory events
        - variability in relative timing
    - language differences

From TADA read me slides

Why TADA: Use example

Parell (2011): Spanish stop spirantization

Observation:

- /b, d, g/ in Spanish - realized as stops only phrase initially, after a homorganic nasal, or, for /d/ only, after /l/.

In all other positions - realized as the voiced approximants [β, ð, ɣ].

- Reliably distinguished in both production and perception (Romero et al. 2007).

Question: Are these stops that spirantize, approximants that undergo fortition, or perhaps something else entirely?

Method: EMA study + simulation study

Why TADA: Use example

Popescu (2019): Temporal organization of liquids in complex syllables

Observation:

- In production, the temporal coordination patterns of laterals and rhotics in coda position differ from those of stops and nasals.
- In perception, words involving coda liquids elicit conflicting syllable count judgments from native speakers.

Question: Unified account for the observed patterns

Method: A cross-linguistic acoustic study + TADA simulations

Why TADA: Use example

Nam et al. (2013): CV combination preferences in babbling

Observation: preference for certain CV combinations in early acquisition across languages

Question: Frame-then-content vs. synergy between tongue movements?

Problem: no study directly measuring the articulators during babbling in order to judge which biomechanical model might be supported. Such data are difficult to obtain, although ultrasound measures are showing promise.

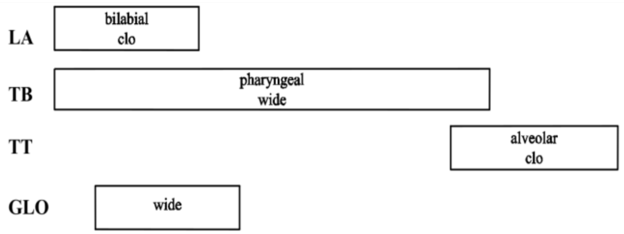
Method: Articulatory synthesis in TADA that allows independent control of tongue, jaw and lips, to model the systematic variation associated with the F/C and the AP accounts.

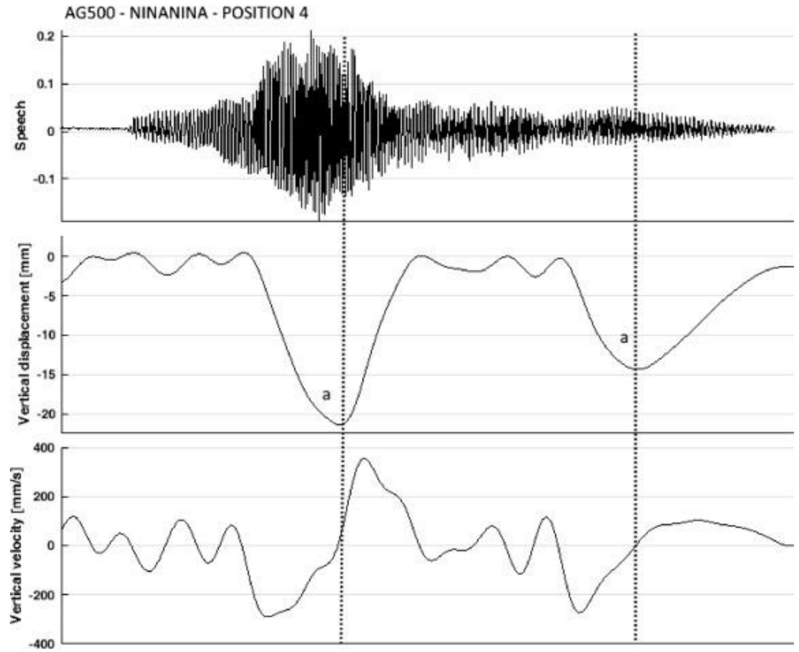
TADA architecture and flow

TADA flow

„model“

TADA

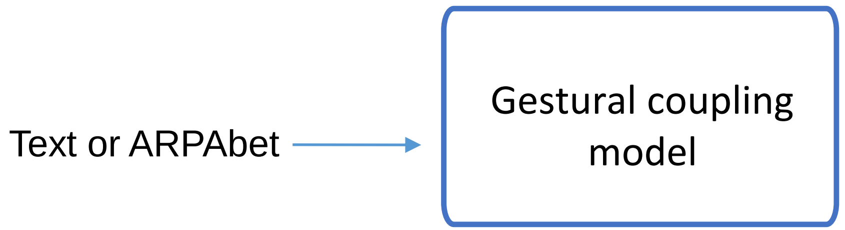




TADA flow

- Open Matlab
- Navigate to the tada folder (or add tada folder to your path)
- Type *´tada‘ 'model' '(MOO)(DAXL)'* ← ARPAbet notation

Name	Date modified	Type	Size
model.HL	17/06/2022 17:46	HL File	10 KB
model	17/06/2022 17:42	MATLAB Data	594 KB
model	17/06/2022 17:42	WAV File	13 KB
model_mv	17/06/2022 17:42	MATLAB Data	107 KB
PHmodel	17/06/2022 17:55	O File	1 KB
TVmodel	17/06/2022 17:55	G File	1 KB
TVmodel	17/06/2022 17:55	O File	2 KB

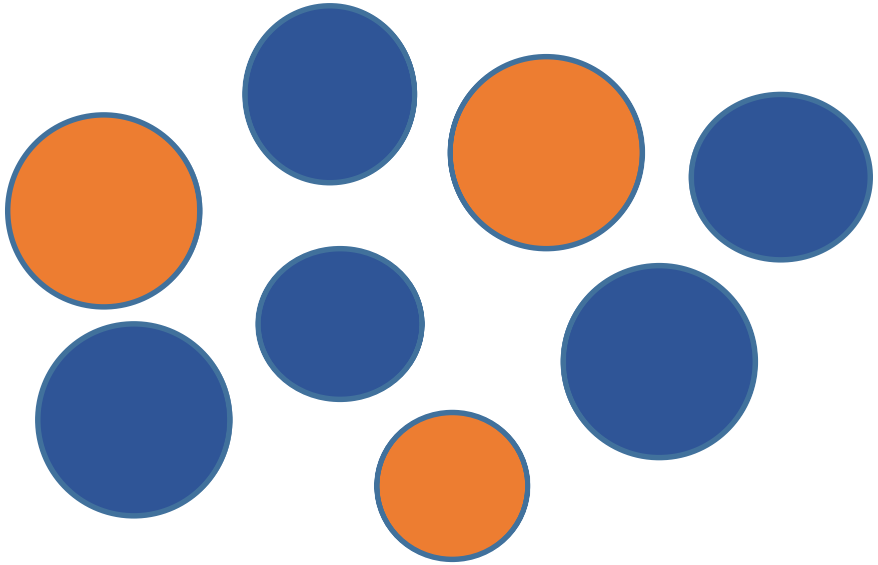

Text or ARPAbet
Gestural coupling model

Articulatory Phonology (Browman & Goldstein):

The basic functional unit of speech is **GESTURE**

Articulatory Phonology (Browman & Goldstein):

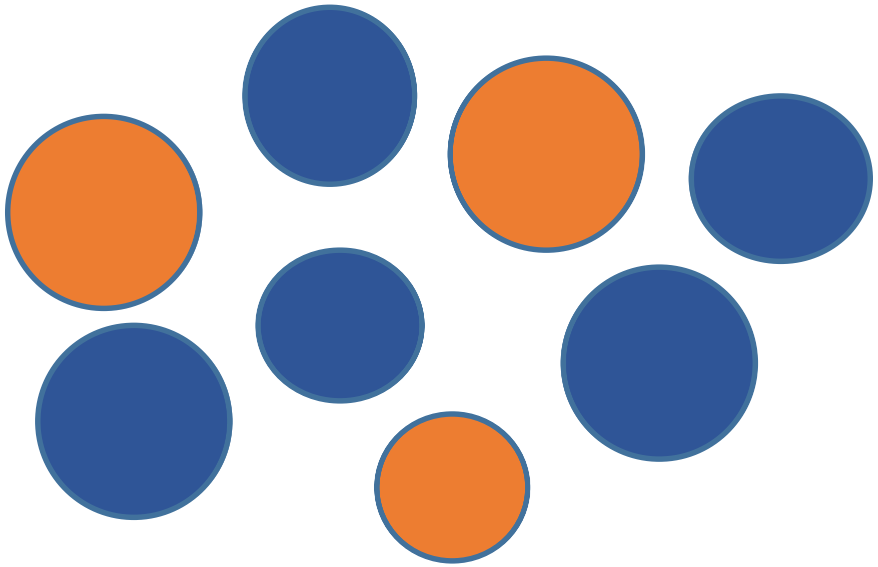
The basic functional unit of speech is **GESTURE**



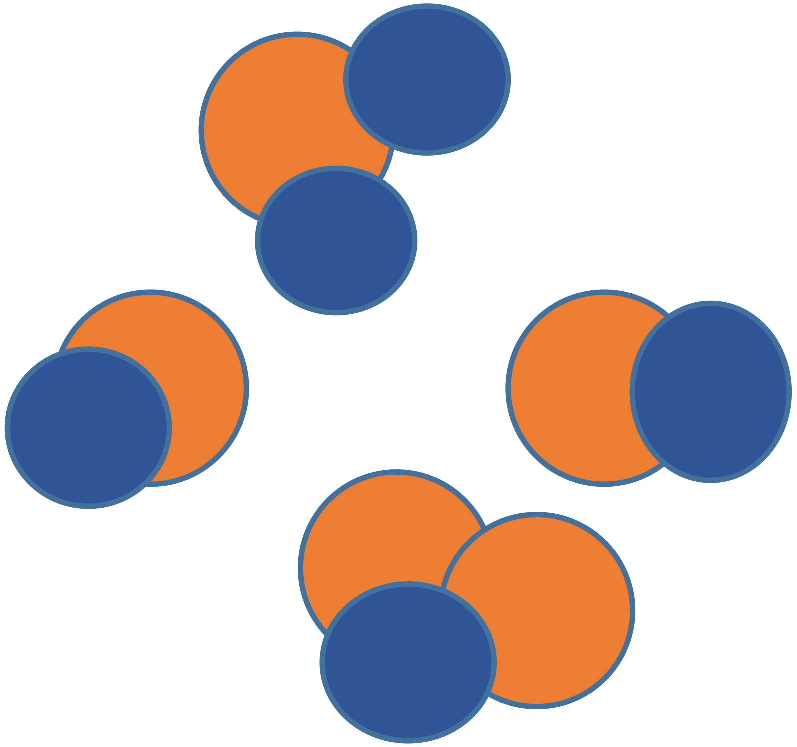
gestures - "atoms"

Articulatory Phonology (Browman & Goldstein):

The basic functional unit of speech is **GESTURE**



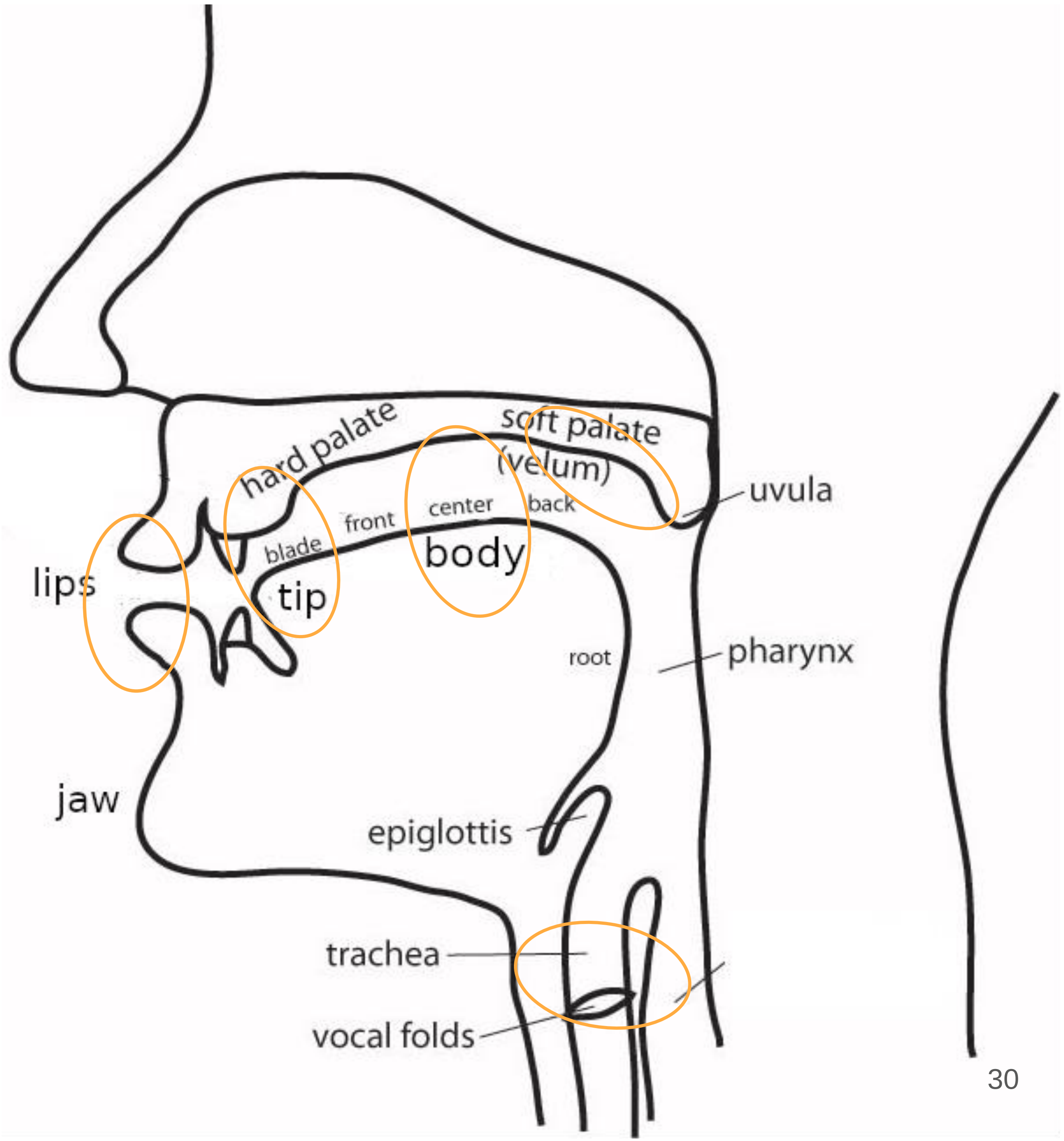
gestures - "atoms"



syllables/words - “molecules”

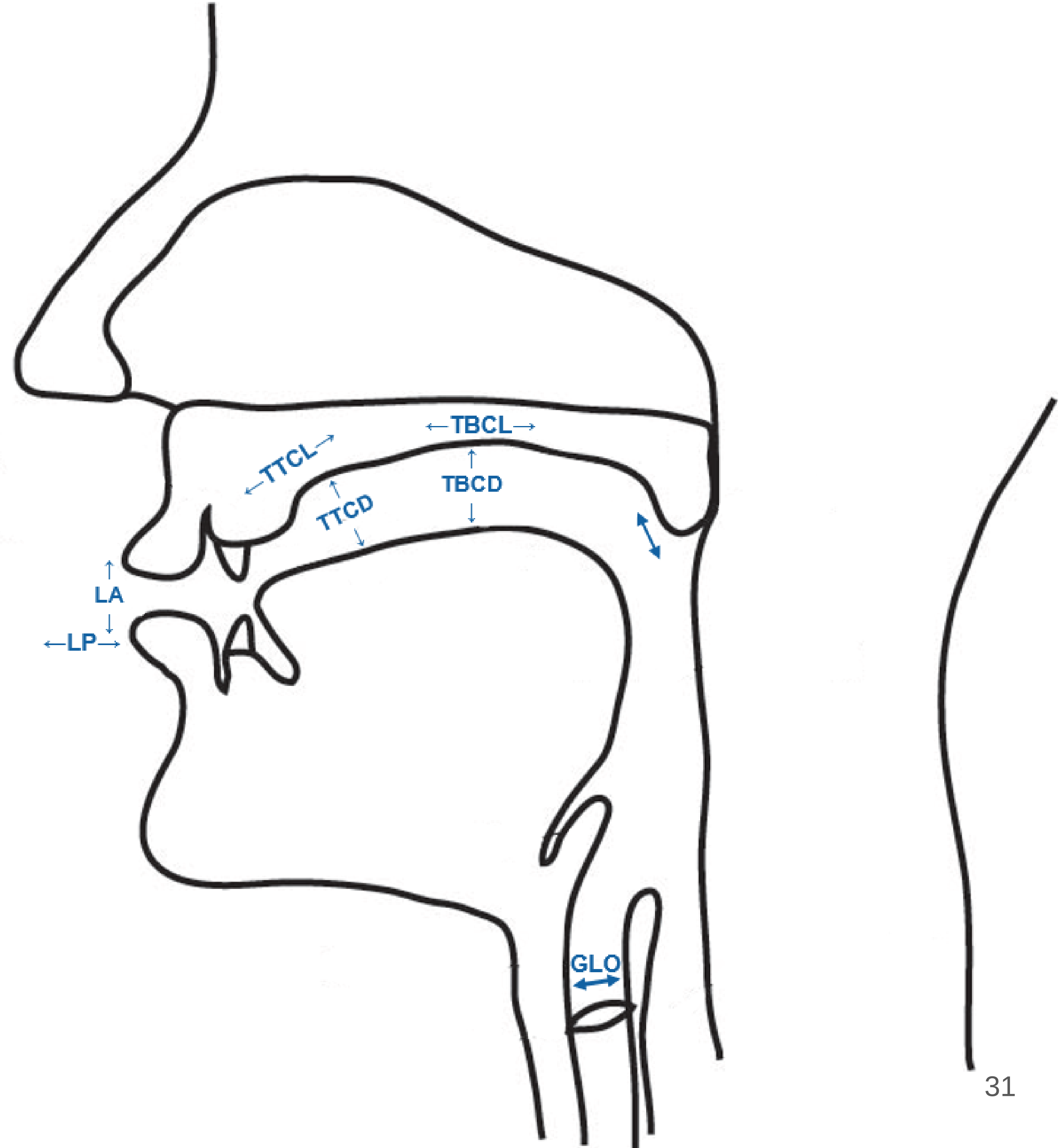
Organs

Lips
Tongue tip
Tongue body
Velum
Glottis

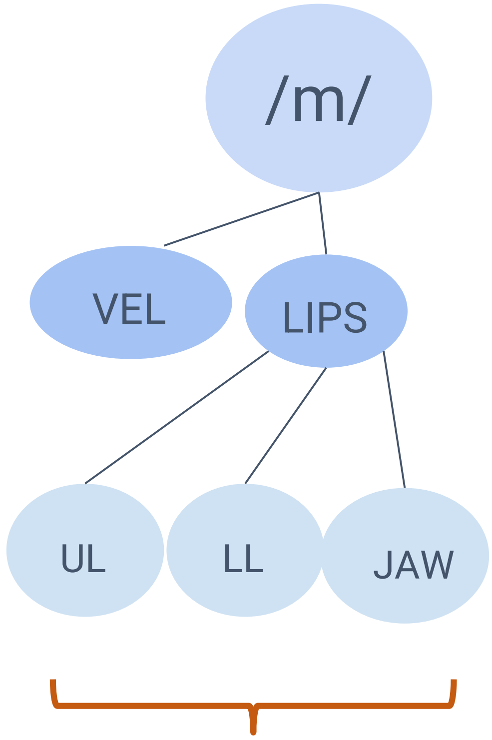


Articulatory Phonology: Tract variables (TVs)

Lip protrusion (LP)
Lip aperture (LA)
Tongue tip constriction location (TTCL)
Tongue tip constriction degree (TTCD)
Tongue body constriction location (TBCL)
Tongue body constriction degree (TBCD)
VEL
GLO



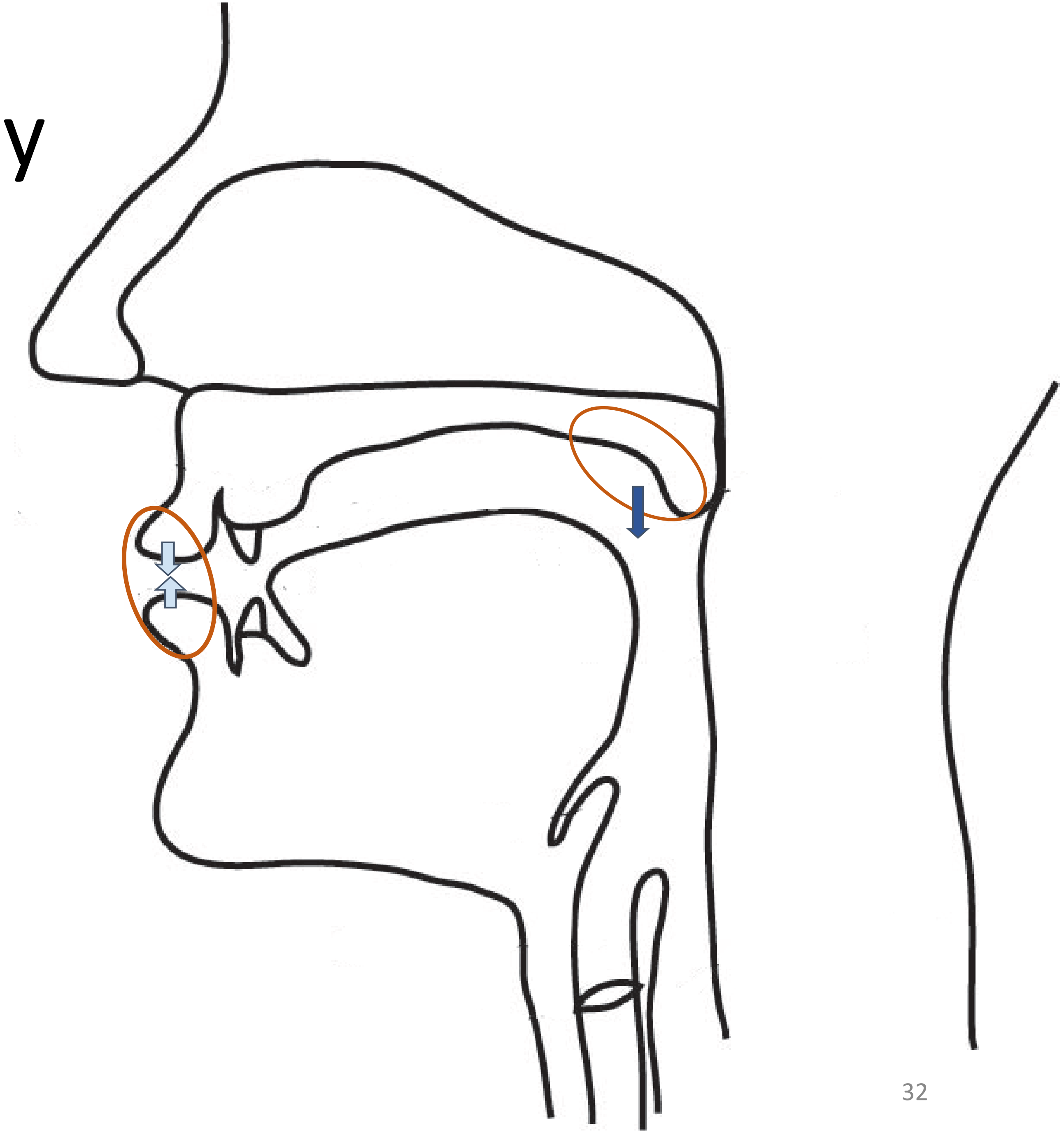
Articulatory Phonology



←gestures

← articulators

synergy
a.k.a coordinative structure
a.k.a control law



Articulatory Phonology

Based on CASY vocal tract model

Constrictors	Tract variables	Articulators involved
Lips	Lip protrusion (PRO)	Upper lip (UH), lower lip (LH), jaw (JA)
	Lip aperture (LA)	Upper lip (UH), lower lip (LH), jaw (JA)
Tongue tip	Tongue tip constriction location (TTCL)	Tongue tip (TL, TA), tongue body (CL, CA), jaw (JA)
	Tongue tip constriction degree (TTCD)	Tongue tip (TL, TA), tongue body (CL, CA), jaw (JA)
Tongue body	Tongue body constriction location (TBCL)	Tongue body (CL, CA), jaw (JA)
	Tongue body constriction degree (TBCD)	Tongue body (CL, CA), jaw (JA)
Velum	Velum (VEL)	Velum (NA)
Glottis	Glottis (GLO)	Glottis (GW)

Articulatory Phonology

Gesture

Unit of information	Unit of action
Phonological	Physical
Linguistic contrast	Control structure
Underlying invariance of representation	Surface variability of perfomance

How to capture this formally?

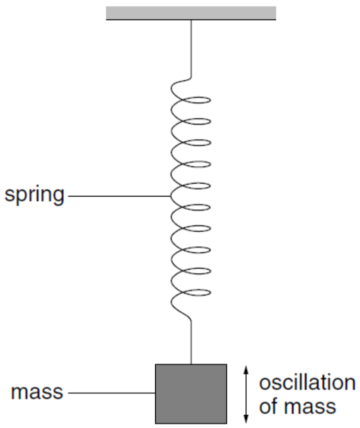
point attractor dynamics of damped oscillator system (i.e., second-order dynamical system)

Gesture as a dynamical system

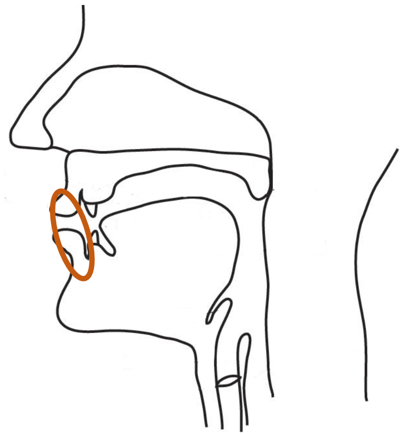
Dynamical system is a system that changes over time according to a set of fixed rules that determine how one state of the system moves to another state.

Examples:

- Position of planets (t)
- voltage of neurons (t,x)
- acoustic waves (t)
- unemployment time series (t,x)
- stock market time series (t)
- the amount of plastic in the ocean (t,x)
- mass-spring system
- position of articulators (t,x)
- and many, many more

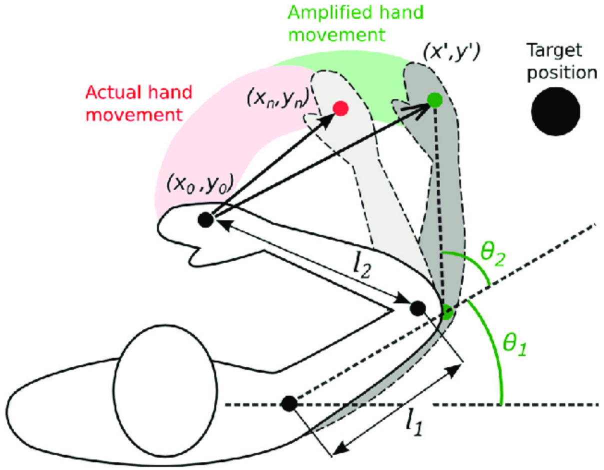


Mass-spring system



Lip Closure

Reaching movement

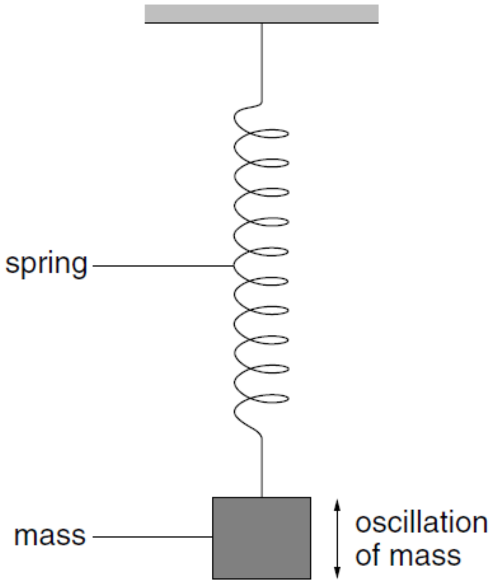


Ballester et al. 2015

Task Dynamics

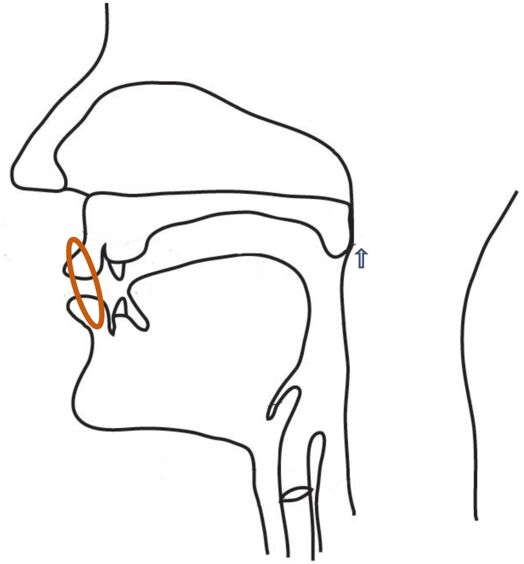
Damped mass-spring system

$$\ddot{x} + \frac{\mu}{m}\dot{x} + \frac{k}{m}x = 0$$



Speech gesture

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



Rule for change:

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

b = damping coefficient

k = stiffness coefficient

x0 = the target position of the tract variable

- Parameter specifications of the system
- Linguistic
- Invariant over time
- Invariant over context

x = current tract-variable positions

x', *x''* = the first and second derivatives of *x* with respect to time

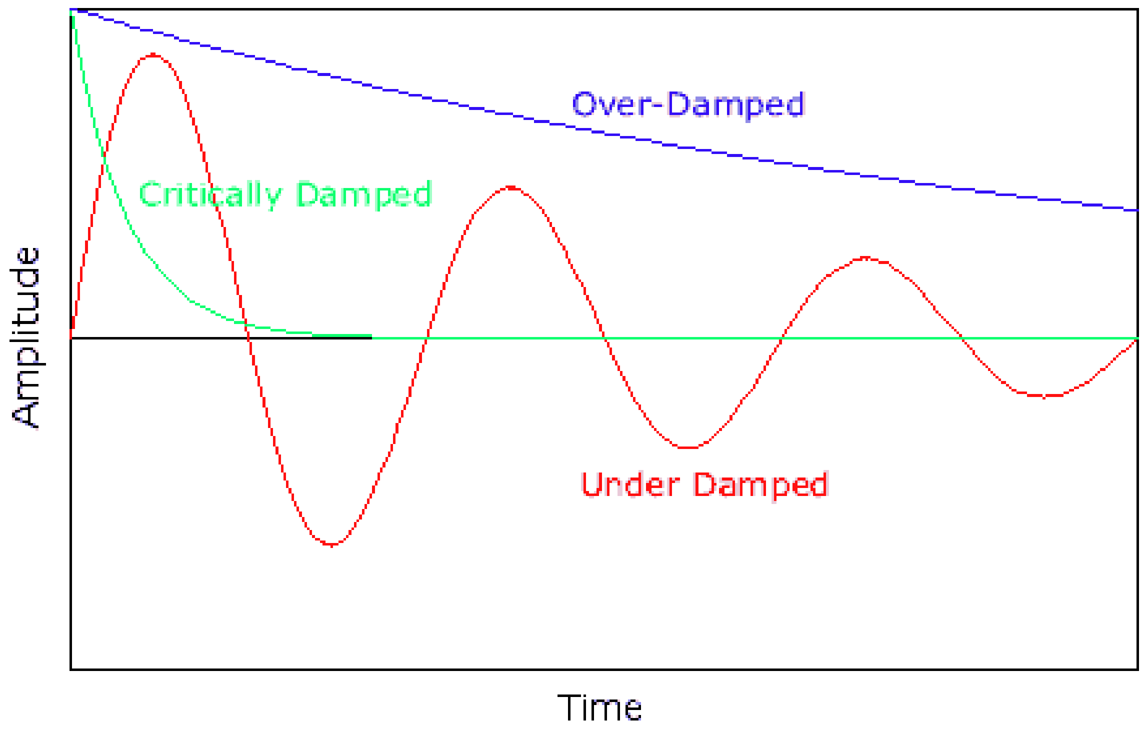
- State variables
- Continuously changing

m = inertial coefficients, *m* is always = 1 (tasks don't have mass)

Rule for change:

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

b = damping coefficient



Rule for change:

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

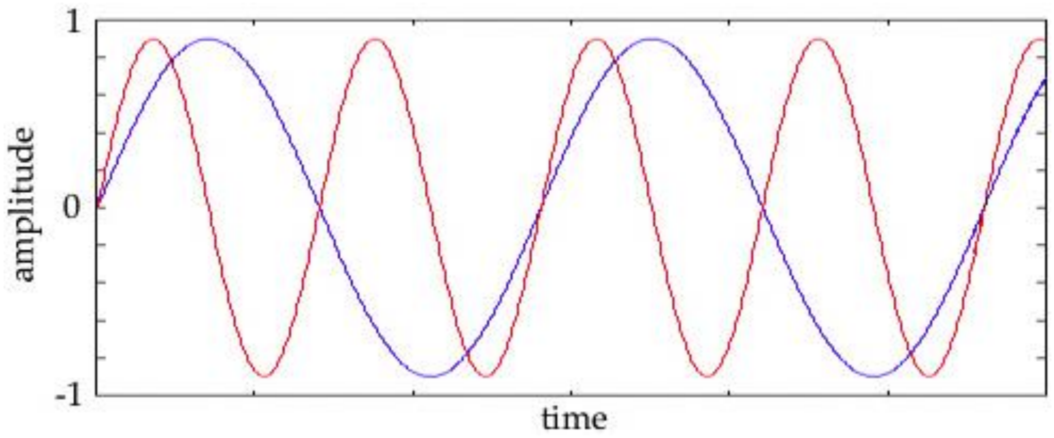
k = stiffness - the *relative* speed with which a gesture reaches its target or moves away from it; specifies at the gestural level, NOT muscular stiffness.

Approximations in American English TADA

For all consonants, k= 8Hz
For all vowels, k = 4Hz
Hz - closure-release cycles per second
-> Consonant:vowel stiffness ratio is 2:1.
(Browman and Goldstein 1990, Nam et al. 2012)

In German TADA:

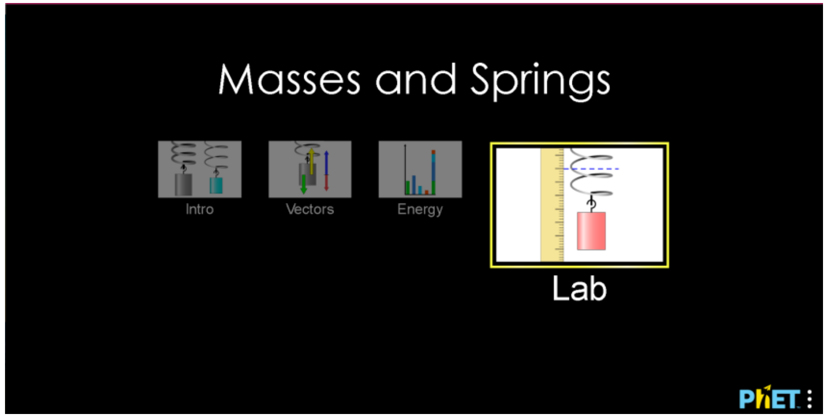
For all vowel TVs, k = 6Hz.
For TT gestures for sonorants and sibilants, k = 10Hz
For TT gestures for stops, k = 12Hz, respectively.
For the consonantal TB target, k = 6Hz
(Pastätter & Pouplier, 2014)



Blue line – lower k
Red line – higher k

Task Dynamics

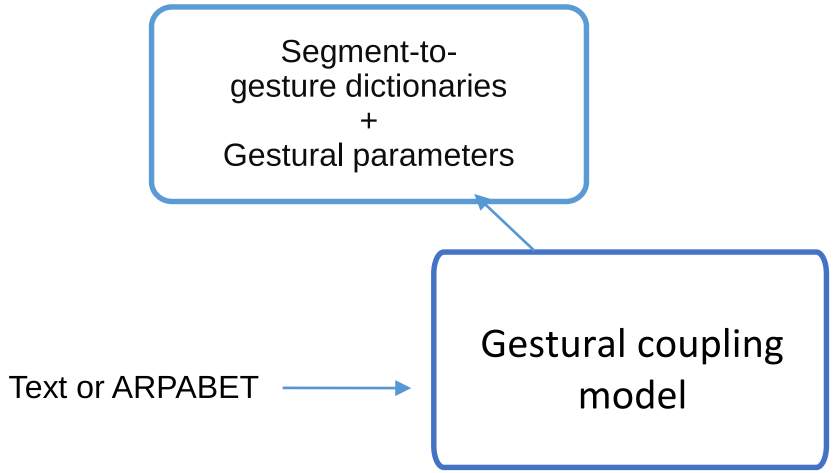
[Mass-spring system simulation]



- Go to Lab
- Change the stiffness (‘spring constant’). Pull the spring. Observe how that changes the velocity of spring movement
- Next, change Damping value from ‘small’ to ‘lots’. Observe how spring oscillation stops after one cycle. That’s our speech gesture - with lots of damping.

Task Dynamics

- Speech motor control - a problem of point attractor dynamics.
- Motor tasks - points in task space, toward which the system is drawn by means of some governing control law which is a function of the system state.
- Control law - damped oscillator system (i.e., second-order dynamical system). Advantages: Damped oscillator dynamics are well-understood, easily characterized, action patterns are globally smooth and continuous.


Segment-to-
gesture dictionaries
+
Gestural parameters
Text or ARPABET
Gestural coupling
model

Gestural parameters

For each tract variable, provides parameters such as target in the vocal tract model, alpha – blending value, list of articulators and their respective weights. The lower the number, the more important articulator for the task.

e.g. For TT gestures JA CL CA TL TA are used The most important ones are TL TA.

TV	Constr	Target	Alpha	LX	JA	UH	LH	CL	CA	TL	TA	NA	GW
TTCL	DENT	40	1	.	32	.	.	32	32	1	1	.	.
TTCL	ALV	56	1	.	32	.	.	32	32	1	1	.	.
TTCL	ALVPAL	60	1	.	32	.	.	32	32	1	1	.	.
TTCL	PAL	80	1	.	32	.	.	32	32	1	1	.	.
TTCL	REL	24	1	.	32	.	.	32	32	1	1	.	.
TTCD	CLO	-2	100	.	32	.	.	32	32	1	1	.	.
TTCD	CRIT	1	10	.	32	.	.	32	32	1	1	.	.
TTCD	NAR	2	1	.	32	.	.	32	32	1	1	.	.
TTCD	REL	11	1	.	32	.	.	32	32	1	1	.	.

gest\gparams.txt

Segment-to-gesture dictionaries

For each phoneme in a given language, gives the list of correspoding geestures: organ, oscillator type, TV, constriction type

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	ARPA	Organ	Osc	TV	Constr	Target	Stiff	Dmp	Alph	LX	JA	UH	LH	CL	CA	TL	TA	NA	GW
2	B	Lips	clo	LA	CLO	.	.	.	.	.	.	.	.	.	.	.	.	.	.
3		Lips	rel	LA	REL	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4		Velum	clo	VEL	CLO	.	.	.	.	.	.	.	.	.	.	.	.	.	.
5	P	Lips	clo	LA	CLO	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6		Lips	rel	LA	REL	.	.	.	.	.	.	.	.	.	.	.	.	.	.
7		Glottis	h	GLO	WIDE	.	16	.	.	.	.	.	.	.	.	.	.	.	.
8		Velum	clo	VEL	CLO	.	.	.	.	.	.	.	.	.	.	.	.	.	.
9	M	Lips	clo	LA	CLO	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10		Lips	rel	LA	REL	.	.	.	.	.	.	.	.	.	.	.	.	.	.
11		Velum	n	VEL	WIDE	.	.	.	.	.	.	.	.	.	.	.	.	.	.
12	D	TT	clo	TTCL	ALV	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13		TT	clo	TTCD	CLO	.	.	.	.	.	.	.	.	.	.	.	.	.	.
14		TT	rel	TTCL	REL	.	.	.	.	.	.	.	.	.	.	.	.	.	.
15		TT	rel	TTCD	REL	.	.	.	.	.	.	.	.	.	.	.	.	.	.
16		Velum	clo	VEL	CLO	.	.	.	.	.	.	.	.	.	.	.	.	.	.
17	T	TT	clo	TTCL	ALV	.	.	.	.	.	.	.	.	.	.	.	.	.	.
18		TT	clo	TTCD	CLO	.	.	.	.	.	.	.	.	.	.	.	.	.	.
19		TT	rel	TTCL	REL	.	.	.	.	.	.	.	.	.	.	.	.	.	.
20		TT	rel	TTCD	REL	.	.	.	.	.	.	.	.	.	.	.	.	.	.
21		Glottis	h	GLO	WIDE	.	16	.	.	.	.	.	.	.	.	.			
22		Velum	clo	VEL	CLO	.	.	.	.	.	.	.	.	.	.	.	.	.	.

gest\seg2gest.txt

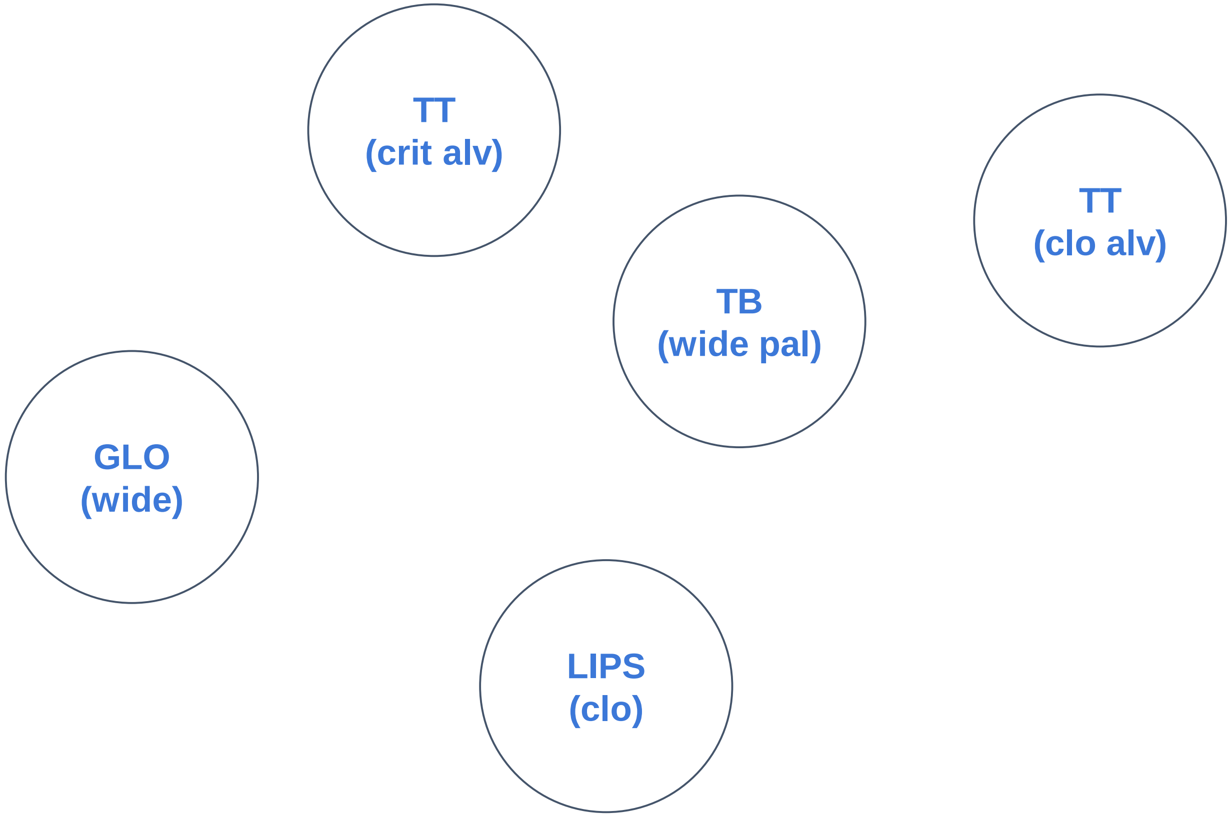
Output: TV<id>.O

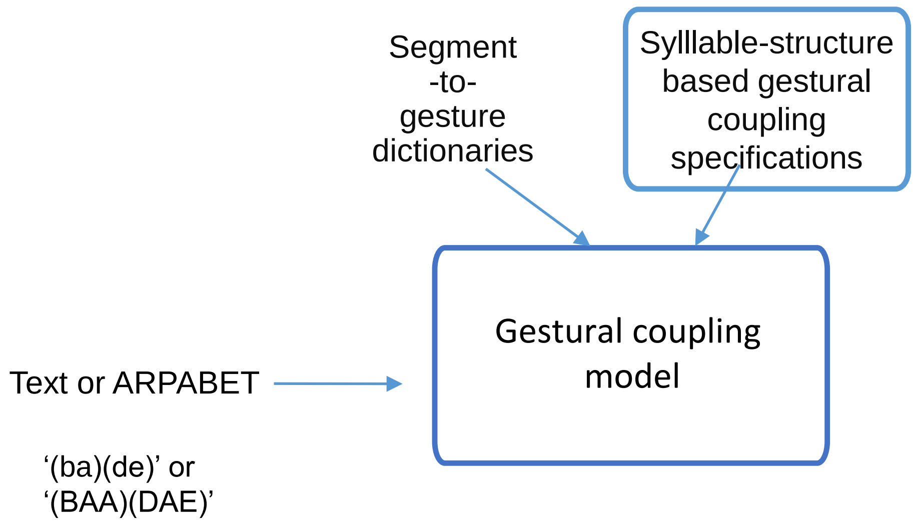
includes

- a list of the constriction gestures in the utterance and for each,
- the values of its control parameters
- a label specifying the oscillator (in the paired PH<id>.o file) that triggers that gesture's activation.

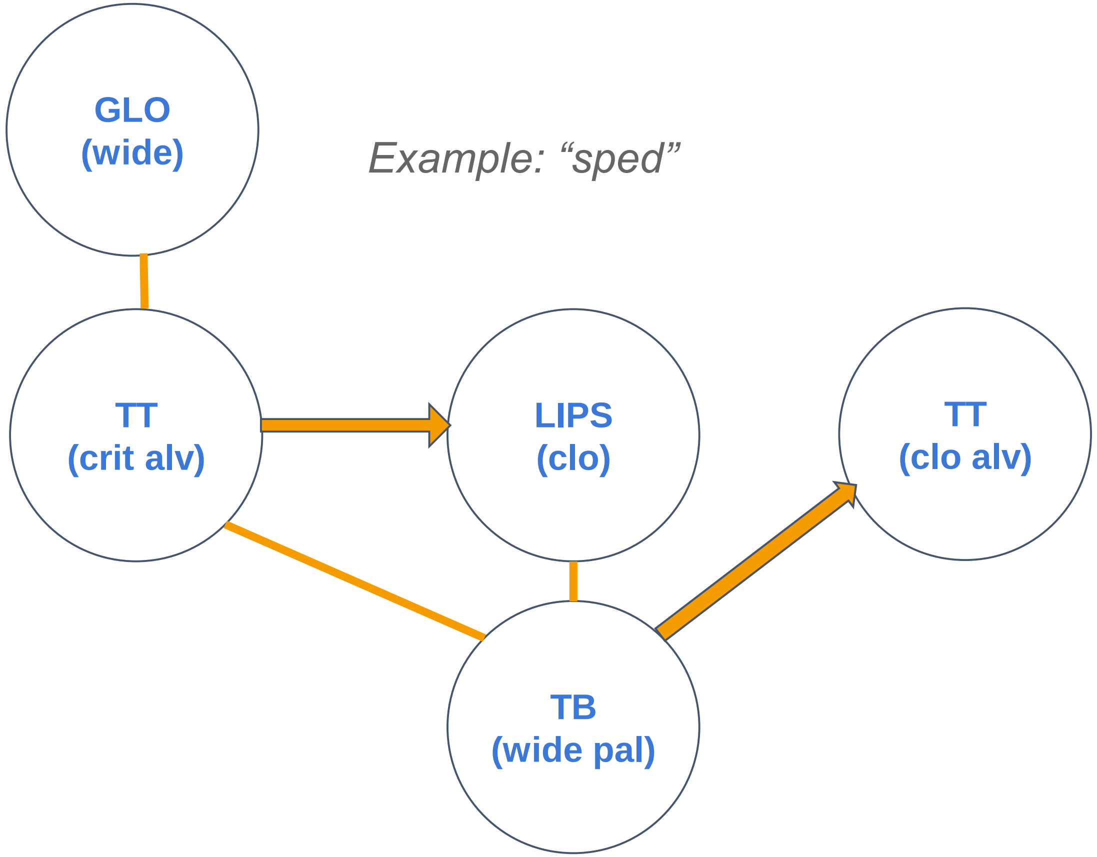
```
% Input string: <(MOO)(DAXL)>
%
%
% Word 1:  (MOO)(DAXL)
% arpabet: (M-OO_)(D-AX_L)
%
%
% syllable 1:   M-OO_
%
%    onset cluster = <M>
%    segment 1 [M]:
'LA' 'ons1_clo1' -2 8 1 JA=8,UH=5,LH=1 100 0.01
'LA' 'ons1_rel1' 11 8 1 JA=8,UH=5,LH=1 1 1
'VEL' 'ons1_n1' 0.2 8 1 NA=1 1 1
%
%    nucleus cluster = <OO>
%    segment 1 [OO]:
'LA' 'v1' 3 6 1 JA=1,UH=5,LH=1 1 1
'TBCL' 'v1' 137 6 1 JA=1,CL=1,CA=1 1 1
'LP' 'v1' 14 6 1 LX=1 1 1
'TBCD' 'v1' 4.70 6 1 JA=1,CL=1,CA=1 1 1
```

Gestures for “sped”




Segment
-to-
gesture
dictionaries
Sylllable-structure
based gestural
coupling
specifications
Gestural coupling
model
Text or ARPABET
‘(ba)(de)’ or
‘(BAA)(DAE)’

Gestural coupling model



Gestural coupling graph -

phonological representation of an utterance

NODES specify gestures

EDGES specify coordination among gestures -

relative phase targets for pairs of gestural oscillators

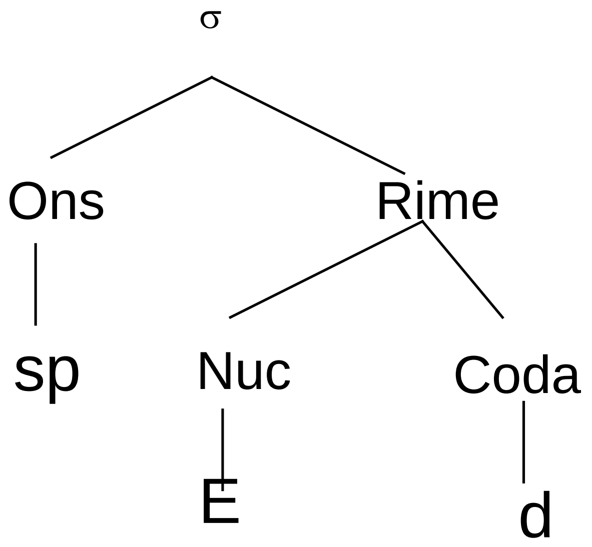
➔ Each gesture is associated with a timing oscillator

➔ Timing oscillators trigger the activation of their associated gesture(s)

LIPS,
TT (Tongue Tip),
TB (Tongue Body),
VEL (Velum),
GLO (Glottis)

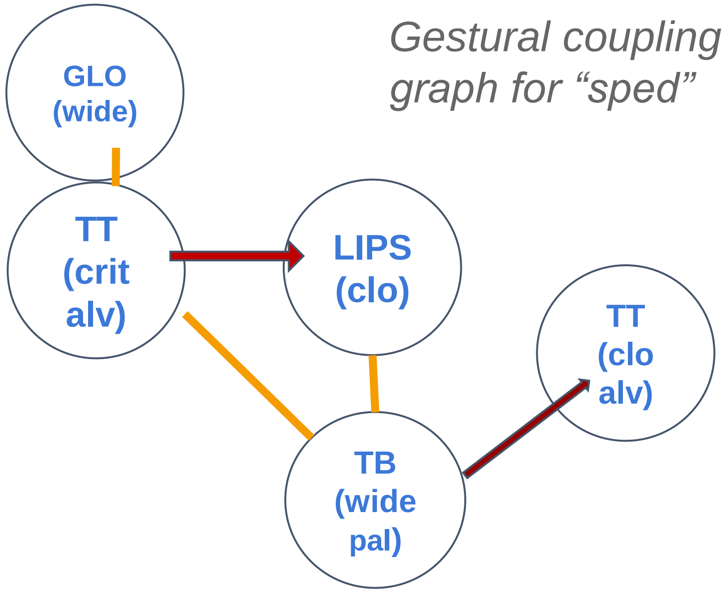
Topology of coupling graph defines syllable structure

Traditional view of syllable



Coupling model of the syllable

- ➔ Onset gestures in-phase to Nucleus gestures
- ➔ Coda gestures anti-phase to Nucleus gestures
- ➔ Oral Constriction gestures within Onset or Coda anti-phase to each other



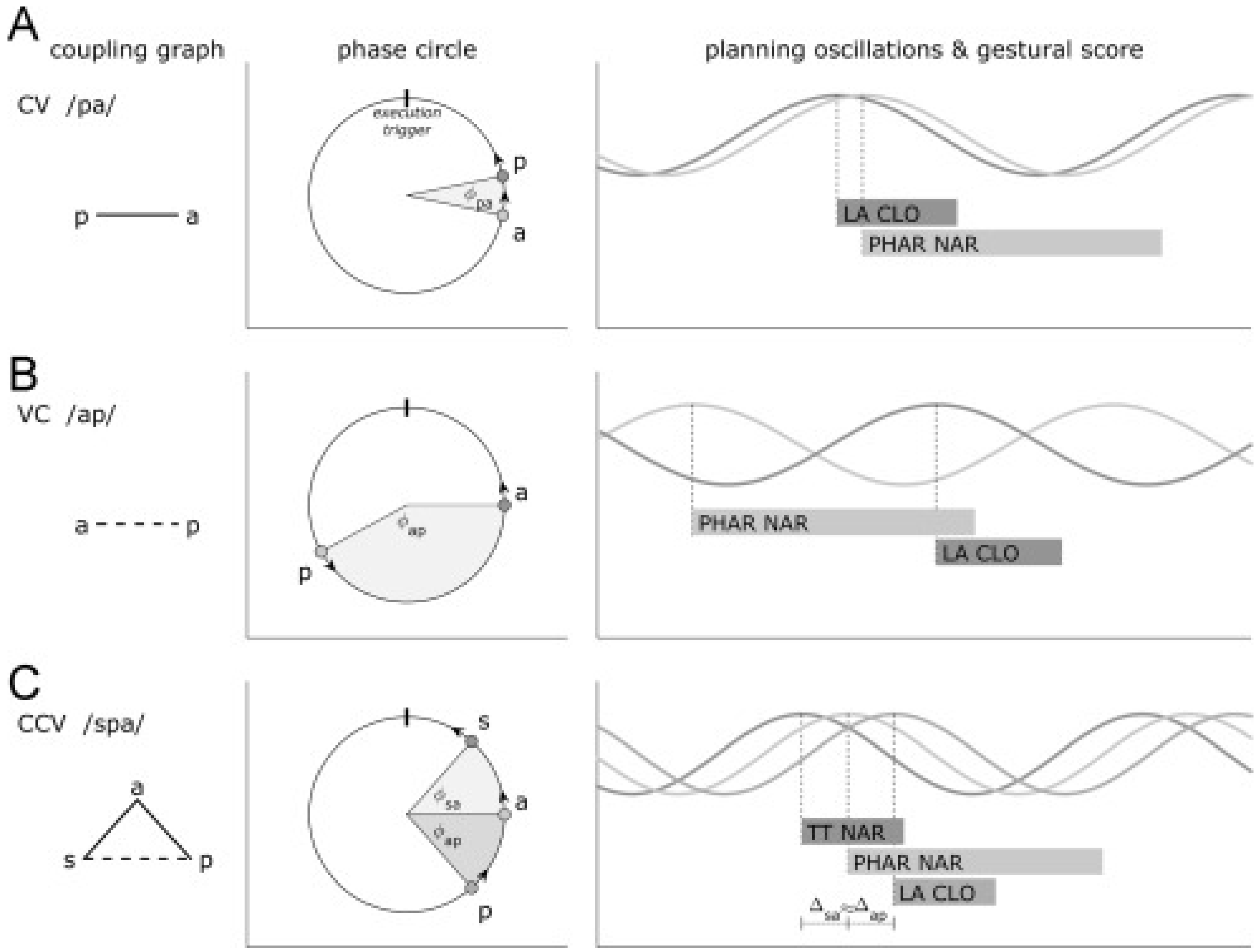
Gestural coupling graph for "sped"

Gestural coupling rules

TADA/gest/coupling.ph

```
%
%     <coupling.ph>
%
%     standard phasing syntax file to generate ph.o file in gest.m
%
%     Created         2009/02/17 Hosung Nam
%     Annotated  2010/04/01 Louis Goldstein
%     Modified   2010/04/02 Michael Proctor (for Perl
implementation)
%     Modified   2012/05/24 Manfred PastÃ¤tter    (rise/ramp/fall
of vowel; lines 13-16); reset 2012/09/13



% oscillatory parameters and activation portion in cycle
v\d+             2 1 4 1 NaN / 10 200 210    % vowel
v_rnd\d+         2 1 4 1 NaN / 10 200 210    % rounding

ons\d*_CLO       2 1 4 1 NaN / 5  60 65      % onset CLO
constriction
ons\d*_REL       2 1 4 1 NaN / 5  20 25      % onset REL
constriction
ons\d*_CRT       2 1 4 1 NaN / 5  60 65      % onset CRT
constriction
ons\d*_NAR       2 1 4 1 NaN / 5  60 65      % onset NAR
constriction
ons\d*_VOC       2 1 4 1 NaN / 5  60 65      % onset VOC
constriction
```

Gestural coupling model



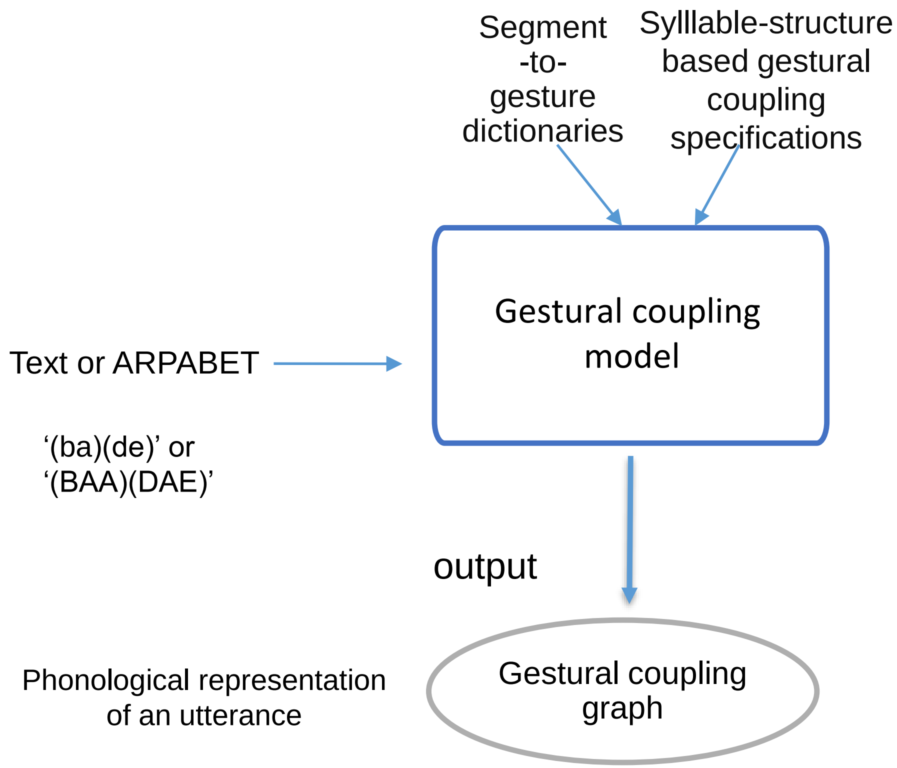
Tilsen (2016)

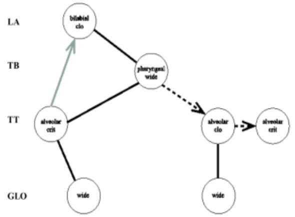
Stages of Speech Production Model

- Planning
  - Gesture oscillators all start at random phases.
  - Over repeated cycles, coupling forces cause oscillators to settle at stabilized relative phases (Saltzman & Byrd, 2000).

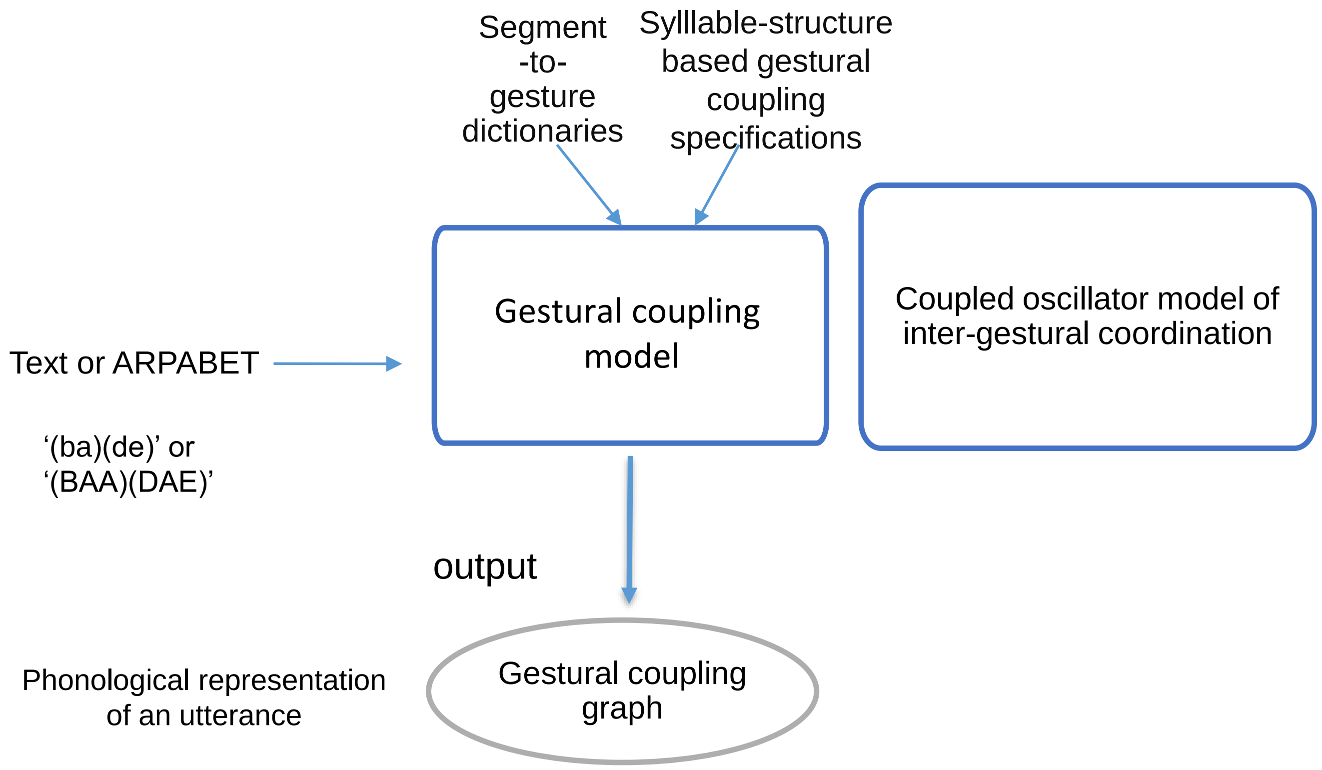
PHmodel.O

```
%'OSC_ID' NatFreq m,n escap amp_init phase_init / riseramp
plateau fallramp
'v1' 2   1   4   1 NaN/ 10 200 210
'v2' 2   1   4   1 NaN/ 10 200 210
'ons1_clo1' 2   1   4   1 NaN/ 5 60 65
'ons1_clo2' 2   1   4   1 NaN/ 5 60 65
'ons1_rel1' 2   1   4   1 NaN/ 5 20 25
'ons1_rel2' 2   1   4   1 NaN/ 5 20 25
'cod2_rel2' 2   1   4   1 NaN/ 5 20 25
'cod2_nar2' 2   1   4   1 NaN/ 5 55 60
'ons1_n1' 2   1   4   1 NaN/ 5 60 65

/coupling/

%'OSC_ID1' 'OSC_ID2' strength1(to OSC1) strength2(to OSC2)
TargetRelPhase
'ons1_clo1' 'ons1_rel1' 1  1 65
'ons1_clo2' 'ons1_rel2' 1  1 65
'ons1_clo1' 'ons1_n1' 1 1 0
'ons1_clo1' 'v1' 1 1 0
'ons1_clo2' 'v2' 1 1 0
'cod2_nar2' 'cod2_rel2' 1  1 60
'v2' 'cod2_nar2' 1   1 180
'v1' 'ons1_clo2' 1   1 180
```





TV<id>.o -gestural specifications
PH<id>.o - timing oscillator and coupling specifications

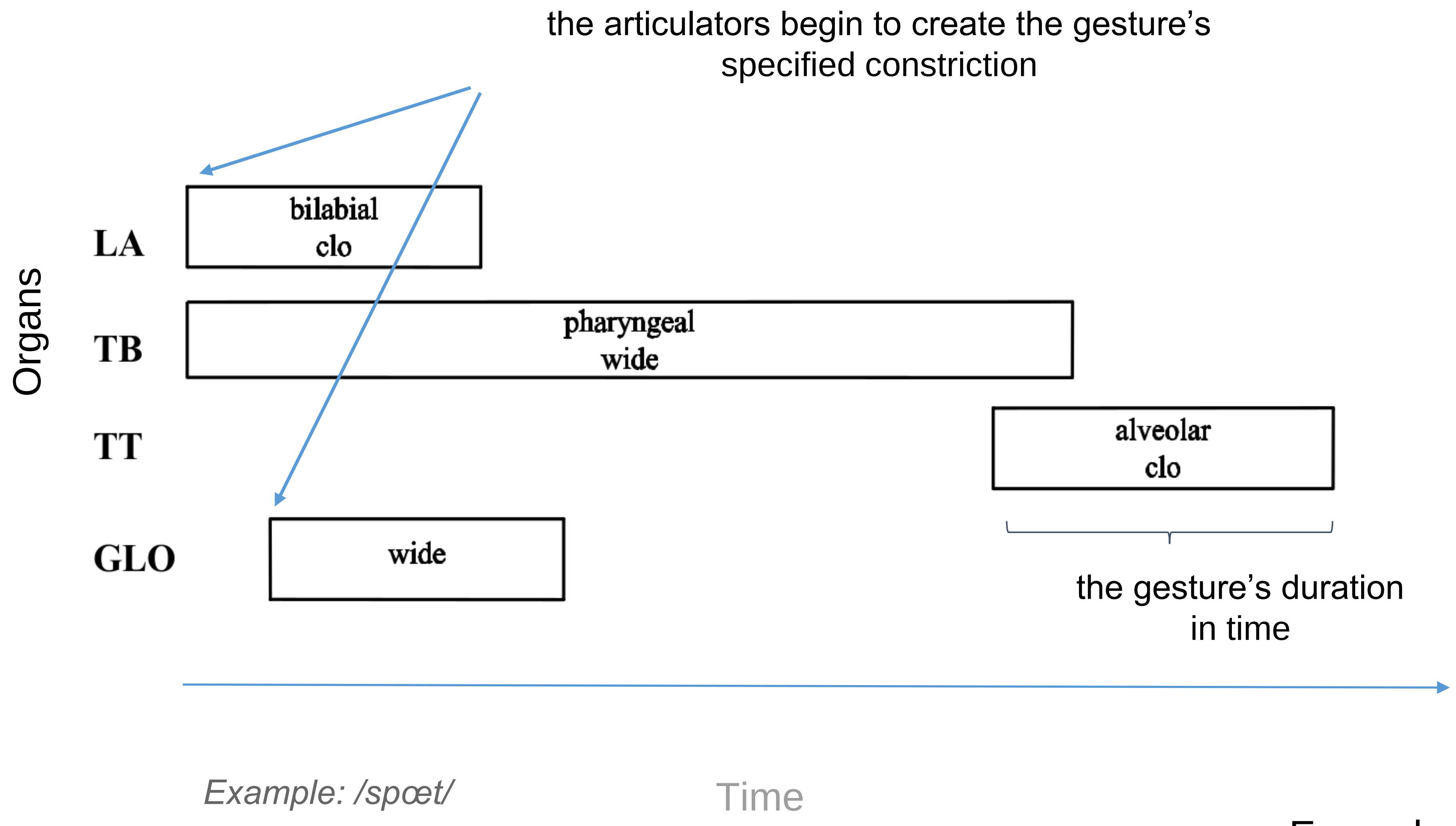


TV<id>.o contains gestural specifications
PH<id>.o contains timing oscillator and coupling specifications

Stages of Speech Production Model

- Planning
  - Gesture oscillators all start at random phases.
  - Over repeated cycles, coupling forces cause oscillators to settle at stabilized relative phases (Saltzman & Byrd, 2000).
  - Cycles of stabilized oscillations used to determine times of gestural activations and deactivations (gestural score)

Gestural score for ‘sped’



Example from Parrell, 2011

Gestural overlap

… may cause conflicting demands on the same articulator.

The blending parameter α (part of a gesture’s specifications) determines the interaction between co-produced gestures.

- α(Cons) = 1 and α(Vowel) = 100 - complete V dominance
- α(Cons) = 1 and α(Vowel) = 1 - an equal blending between C and V

- Blending parameters need to be language-specific (Iskarous et al., 2012)

Gestural overlap

Example: /k/ and vowels

- both produced using the same task variables (TBCL, TBCD)

    - TBCL: α(/k/) = 10 and α(Vowel) = 1
    - - slight C dominance, TBCL of /k/ is shifted in a back vowel context.

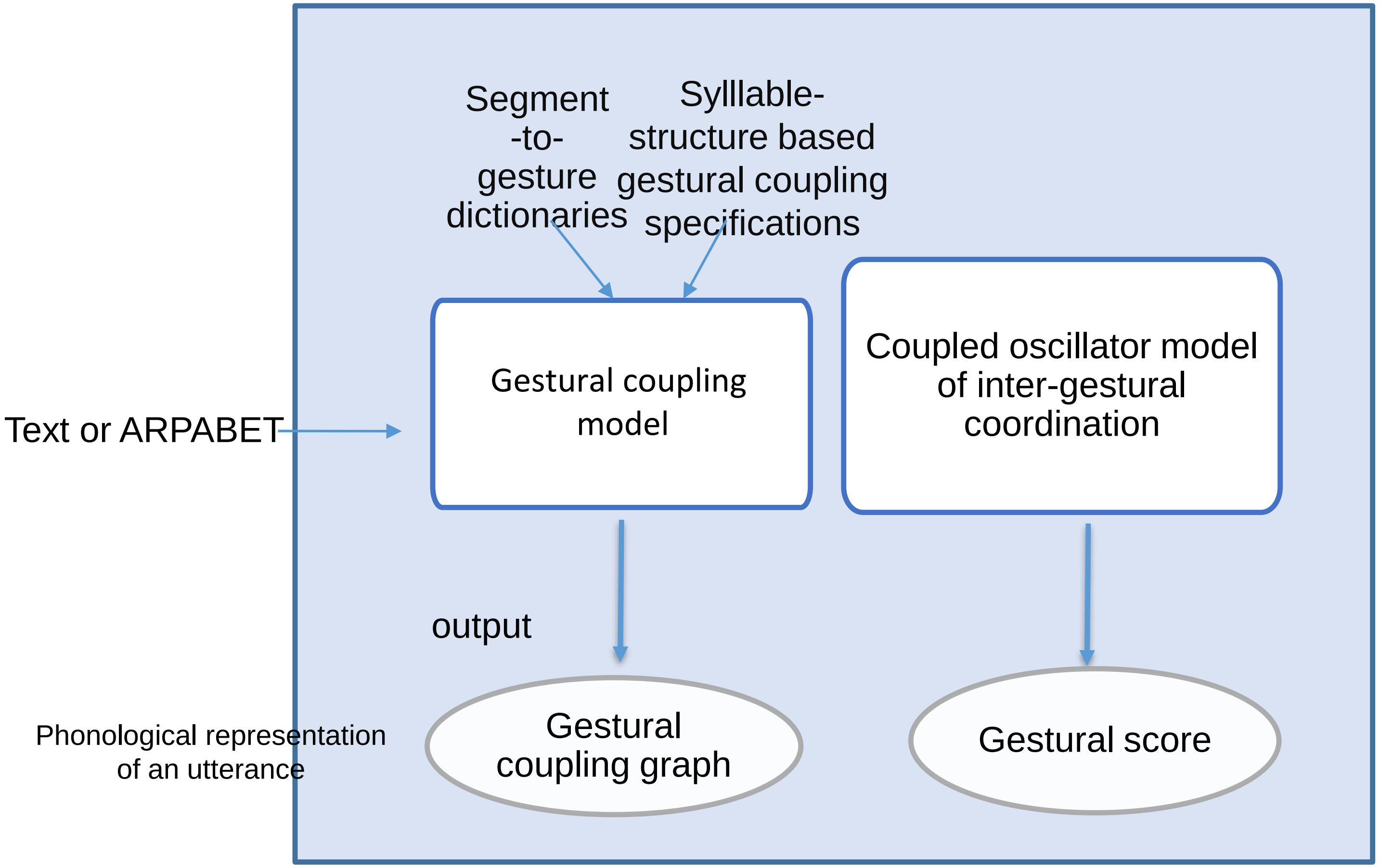
    - TBCD: α(/k/) = 100 and α(Vowel) = 1
    - - complete C dominance, supresses CD of vowel so complete /k/ closure is achieved

/k/ in "skills"

/k/ in "cutbacks"

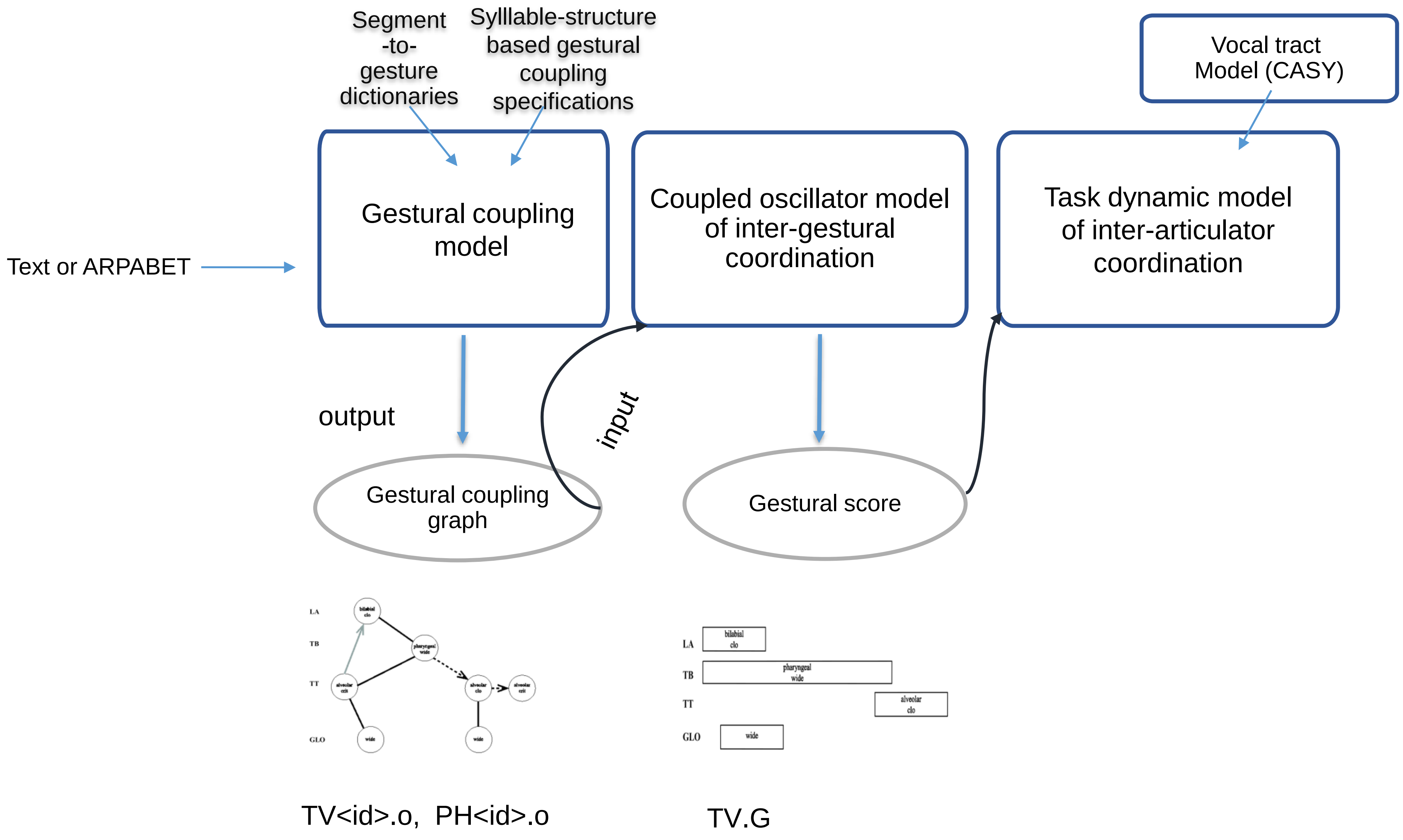
Goldstein, 2022

Planning



TV<id>.o contains gestural specifications
PH<id>.o contains timing oscillator and coupling specifications

TV.G


Segment
-to-
gesture
dictionaries
Sylllable-structure
based gestural
coupling
specifications
Vocal tract
Model (CASY)
Text or ARPABET
Gestural coupling
model
Coupled oscillator model
of inter-gestural
coordination
Task dynamic model
of inter-articulator
coordination
output
input
Gestural coupling
graph
Gestural score
LA
TB
TT
GLO
bilabial
clo
pharyngeal
wide
alveolar
clo
wide
TV<id>.o, PH<id>.o
TV.G

Task-dynamic model

Phonological task equations for each active task (forward dynamics).

$$\ddot{z} = M^{-1}(-B\dot{z} - K\Delta z),$$

z is a vector of Task Variables,
M, B, K are corresponding task parameters
(mass, damping, stiffness)

The tract variables (z) are transformed to model articulators as follows

$$z = z(\phi)$$

$$\dot{z} = J(\phi)\dot{\phi}$$

$$\ddot{z} = J(\phi)\ddot{\phi} + \dot{J}(\phi,\dot{\phi})\dot{\phi},$$

Φ - a vector of Articulator postures [Φ1 Φ2 ... Φ10]
z - a vector of task variables
J, Jacobian - the partial derivative of each task variable with respect to each articulator

Φ and J can be calculated analytically from the geometry of the simplified articulatory model (CASY)

Goldstein, 2022

MODEL ARTICULATOR VARIABLES

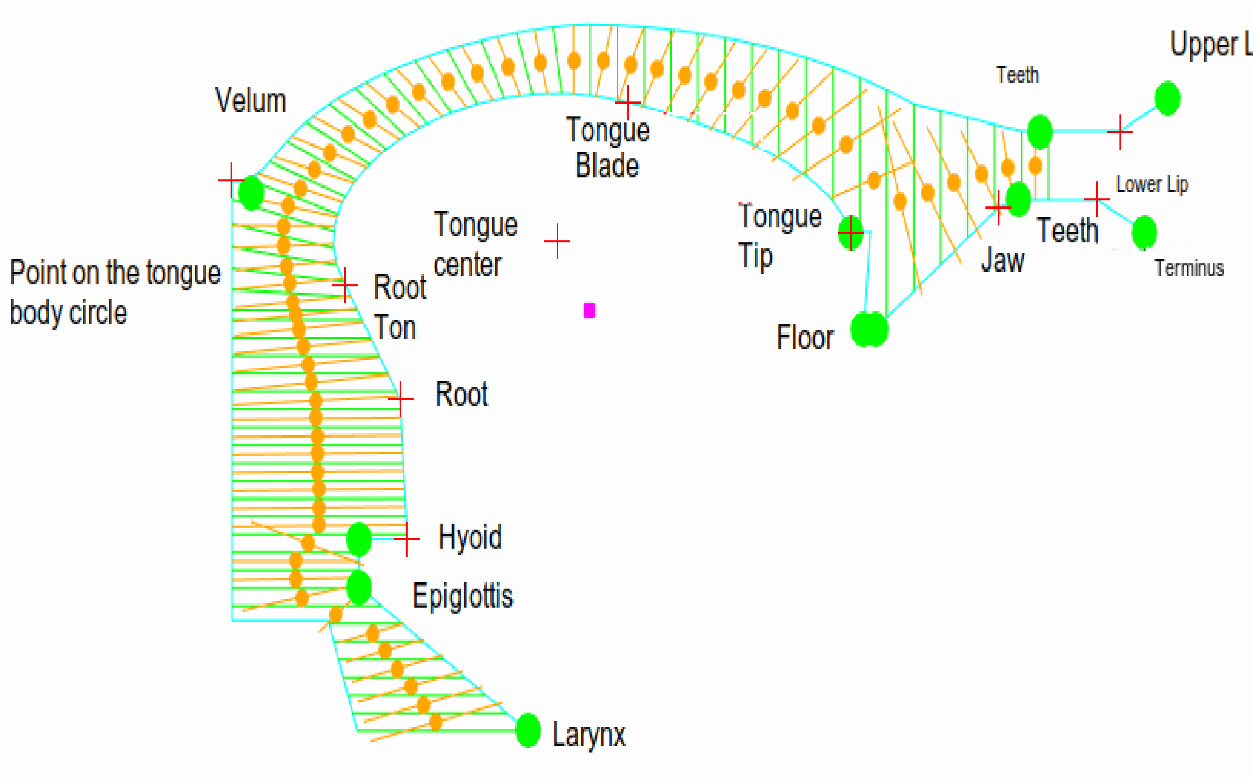
($\varnothing_j$; j = 1, 2, ...,n; n=10)

TRACT VARIABLES (Z_i ; i = 1, 2, ...,m; m=9)	LH ($\varnothing_1$)	JA ($\varnothing_2$)	ULV ($\varnothing_3$)	LLV ($\varnothing_4$)	TBR ($\varnothing_5$)	TBA ($\varnothing_6$)	TTR ($\varnothing_7$)	TTA ($\varnothing_8$)	V ($\varnothing_9$)	G ($\varnothing_{10}$)
LP (Z_1)	●									
LA (Z_2)		●	●	●						
TDCL (Z_3)		●			●	●				
TDCD (Z_4)		●			●	●				
LTH (Z_5)		●								
TTCL (Z_6)		●			●	●	●	●		
TTCD (Z_7)		●			●	●	●	●		
VEL (Z_8)									●	
GLO (Z_9)										●

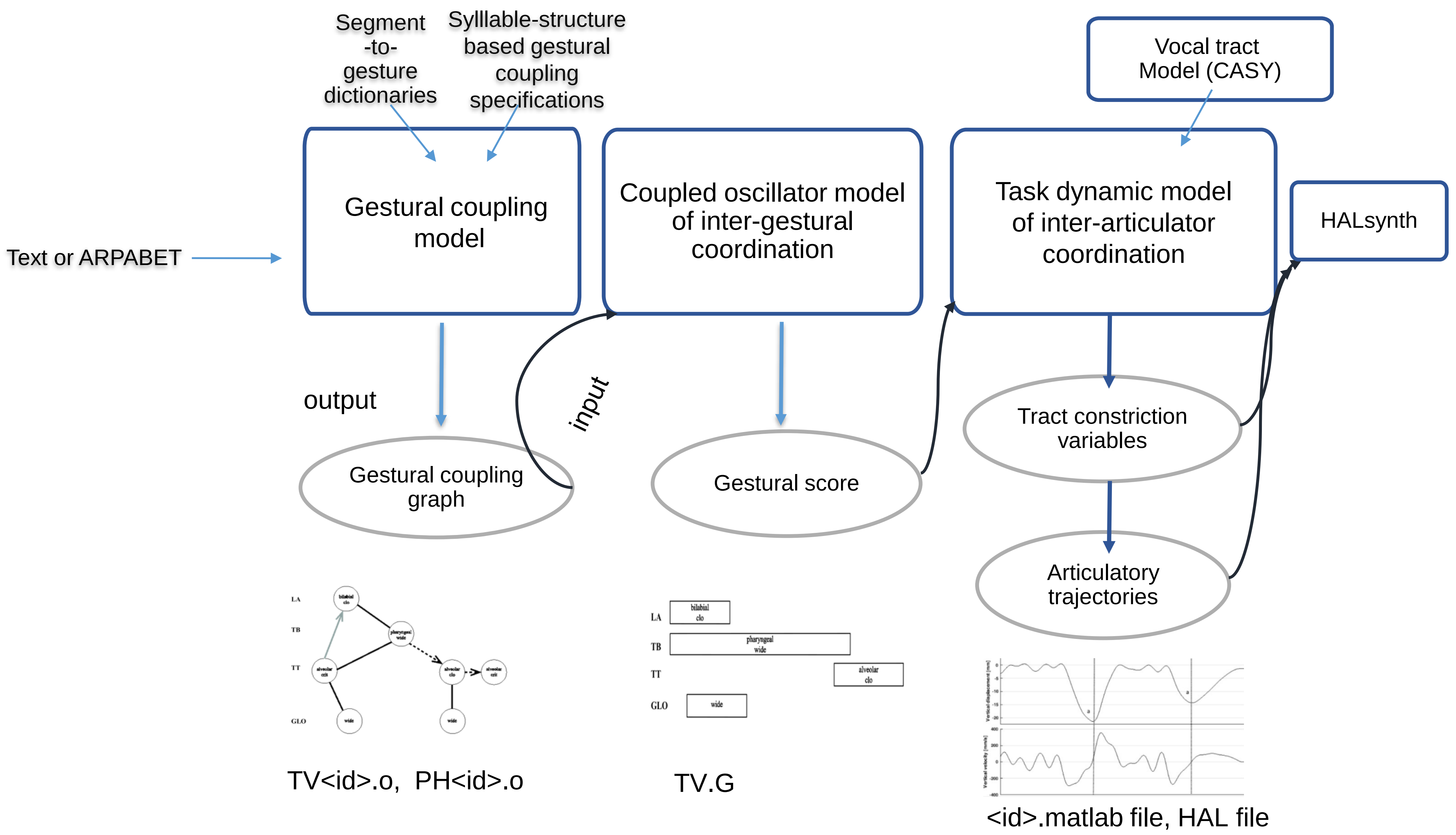
FIGURE 4 Matrix representing the relationship between tract-variables (z) and model articulators (ϕ). The filled cells in a given tract-variable row denote the model articulator components of that tract-variable's articulatory set. The empty cells indicate that the corresponding articulators do not contribute to the tract-variable's motion. (See text for definitions of abbreviations used in the figure.)

CASY Vocal Tract Model

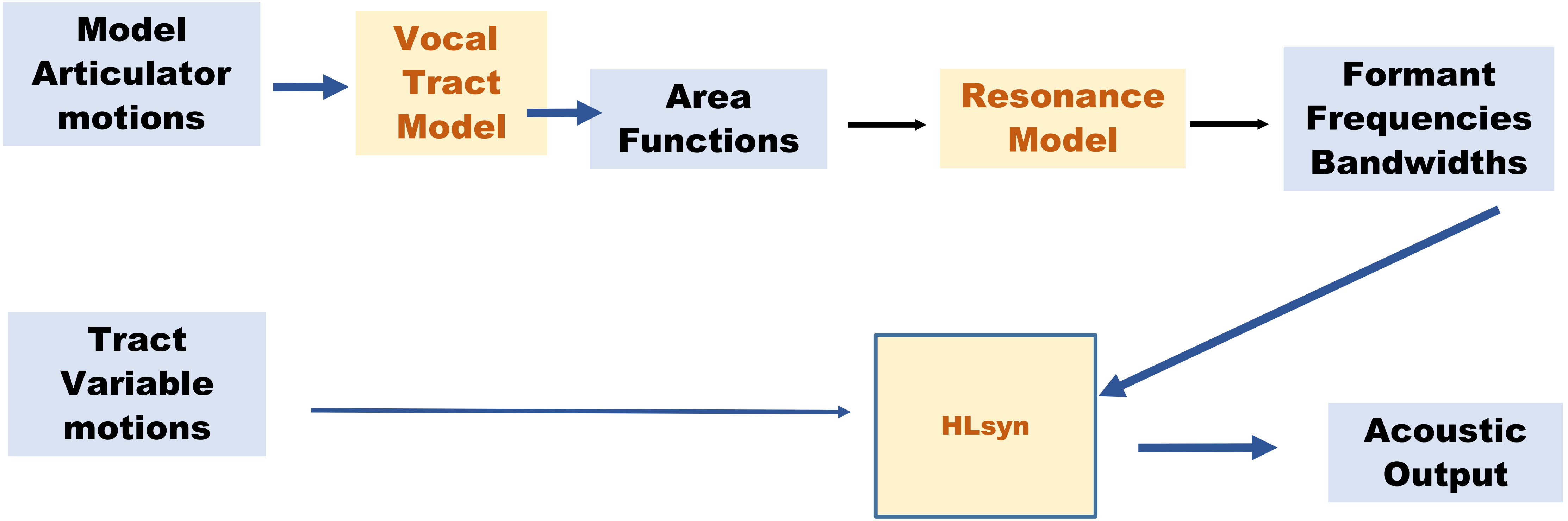
- a geometric model of the vocal tract, similar in spirit to Maeda's model.



Lingustic gestural model

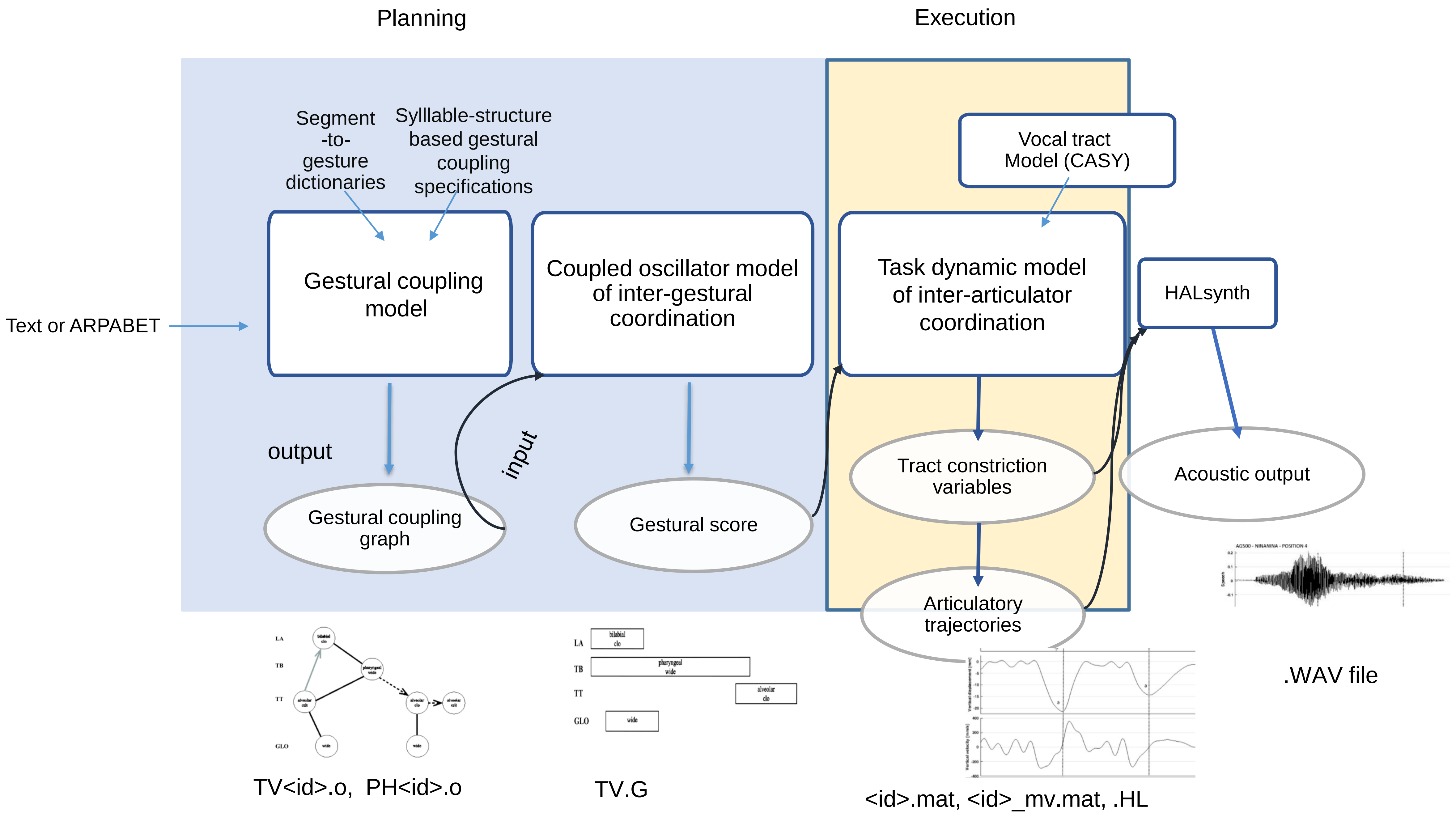


Generation of acoustic output



TADA outputs

- Model articulator time functions
- Constriction (Tract Variable) time functions
- Pseudo-sound
- Input file to HLsyn *(<id>.HL*)



Planning
Execution
Segment -to- gesture dictionaries
Sylllable-structure based gestural coupling specifications
Text or ARPABET
Gestural coupling model
Coupled oscillator model of inter-gestural coordination
Vocal tract Model (CASY)
Task dynamic model of inter-articulator coordination
HALsynth
output
input
Gestural coupling graph
Gestural score
Tract constriction variables
Acoustic output
Articulatory trajectories
LA
TB
TT
GLO
bilabial clo
pharyngeal wide
alveolar clo
wide
TV<id>.o, PH<id>.o
TV.G
<id>.mat, <id>_mv.mat, .HL
.WAV file

Prosody

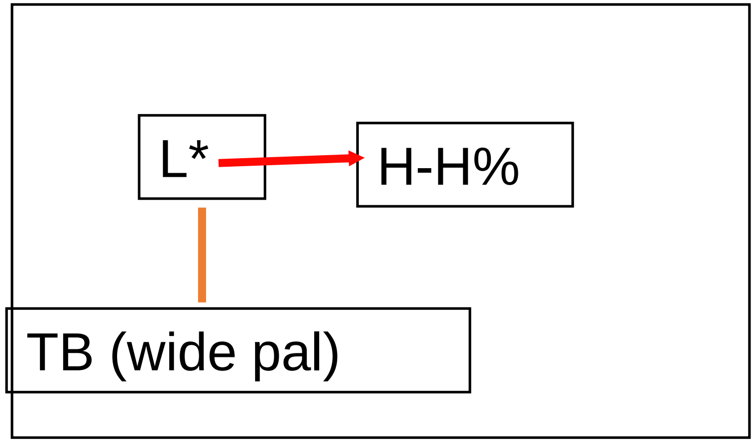
- No prosodic structure or stress is automatically generated at present
- However, following prosodic gestures can be added manually
  - Tone gestures
  - Π-gestures

Prosody

Tone gestures (goal = f0)

- can be added to TV.O file manually
- can be coupled to other gestures

Example: “Cathi Best?”



Prosody

Π-gestures (Byrd & Saltzman, 2002)

- Act to slow the gestural activation clock, effectively stretching gestures in time
- Activation intervals for Π-gestures can be added manually to gestural score.
- Effects: gestural lengthening, spatial strengthening, reduced intergestural overlap

Example: “Cathi Best?” with Π-gesture added.

TADA: Limitations

- No auditory feedback – e.g. no simulations of auditory feedback experiments (for auditory feedback, see SFC & FACTS)
- Vocal tract model not biomechanical
- Some prosodic effects are not explained
- No learning

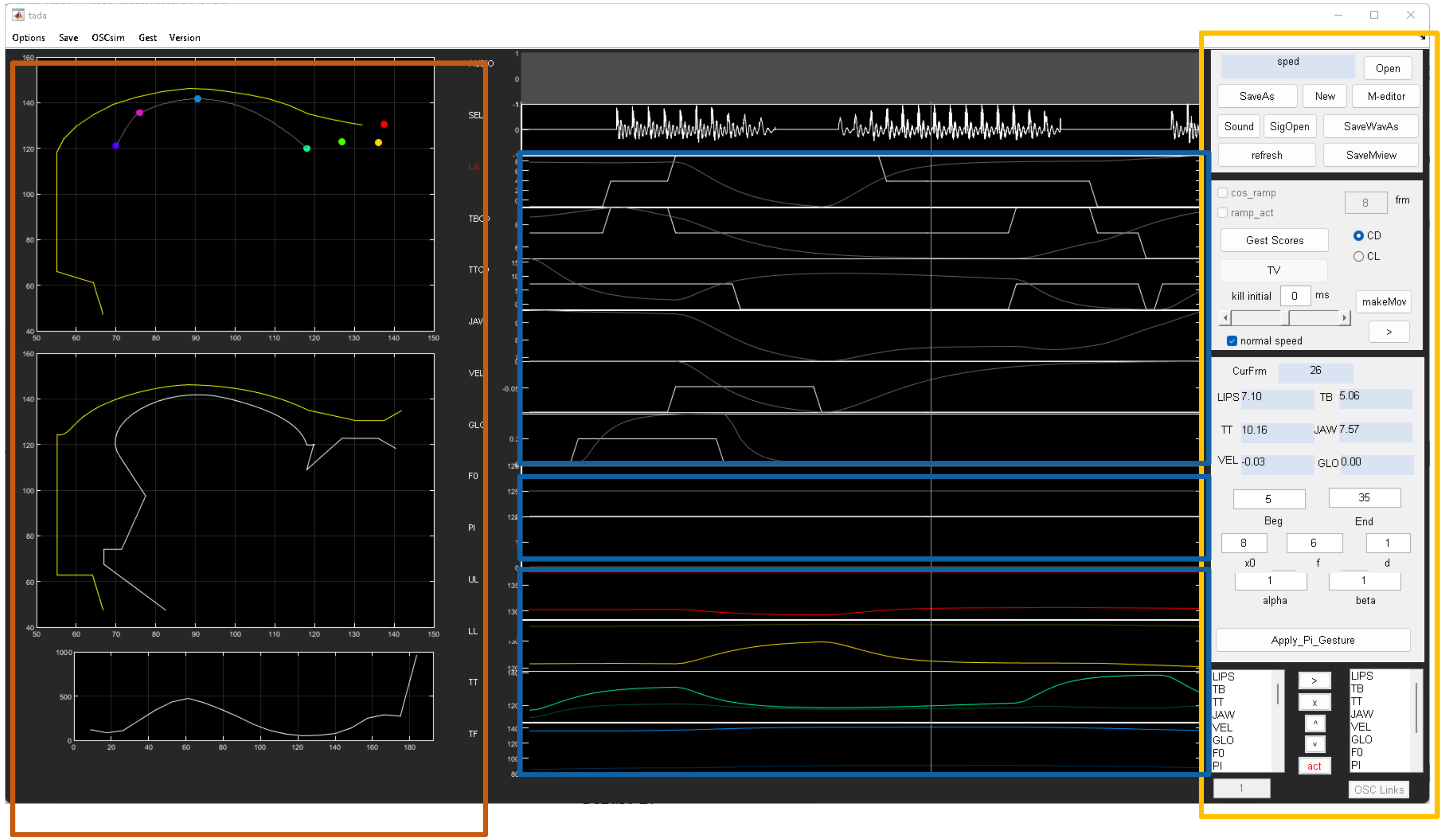
TADA GUI

- LA gesture for /p/ in /sped/ selected

temporal display
from top:
- Gestural score and time functions of tract variables
- Model articulator trajectories

spatial display
at the time of the cursor in the temporal display

from top:
- Model articulator positions (TR, TD, TF, TT, UL, LL, JA)
- Vocal tract shape
- Area function



from top:
1. File manipulation buttons
2. Program action controls
3. Numerical readouts (and one Action button)
4. Time function layout editor

Editing Capabilites in TADA GUI

- Gestural score editing in TADA GUI
  - Graphical
    - Sliding gestures (or groups) in time
    - Stretching or shrinking activation intervals
    - Deletion of gestures
    - Addition of gestures
  - Numerical
    - Dynamical parameters
    - Blending parameters

Simulation experiment

Coarticulation across development

Simulation experiment

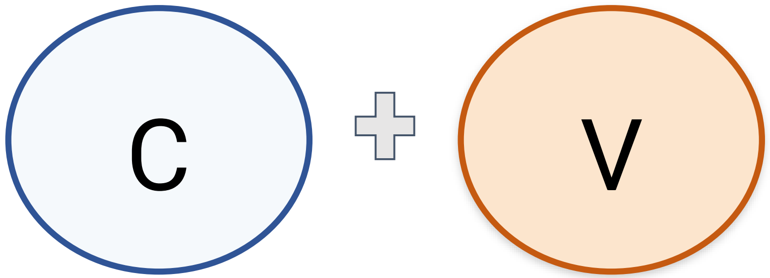
- a test or a series of tests in which meaningful changes are made to the input variables of a simulation model so that we may observe and identify the reasons for changes in the performance measures.

Steps:

1. Formulate a research question that a model and simulation can answer
2. Prototype your methods and create a verification and validation plan
3. Obtain experimental data
4. Calibrate
5. Validate your results
6. Document and share your model and simulation

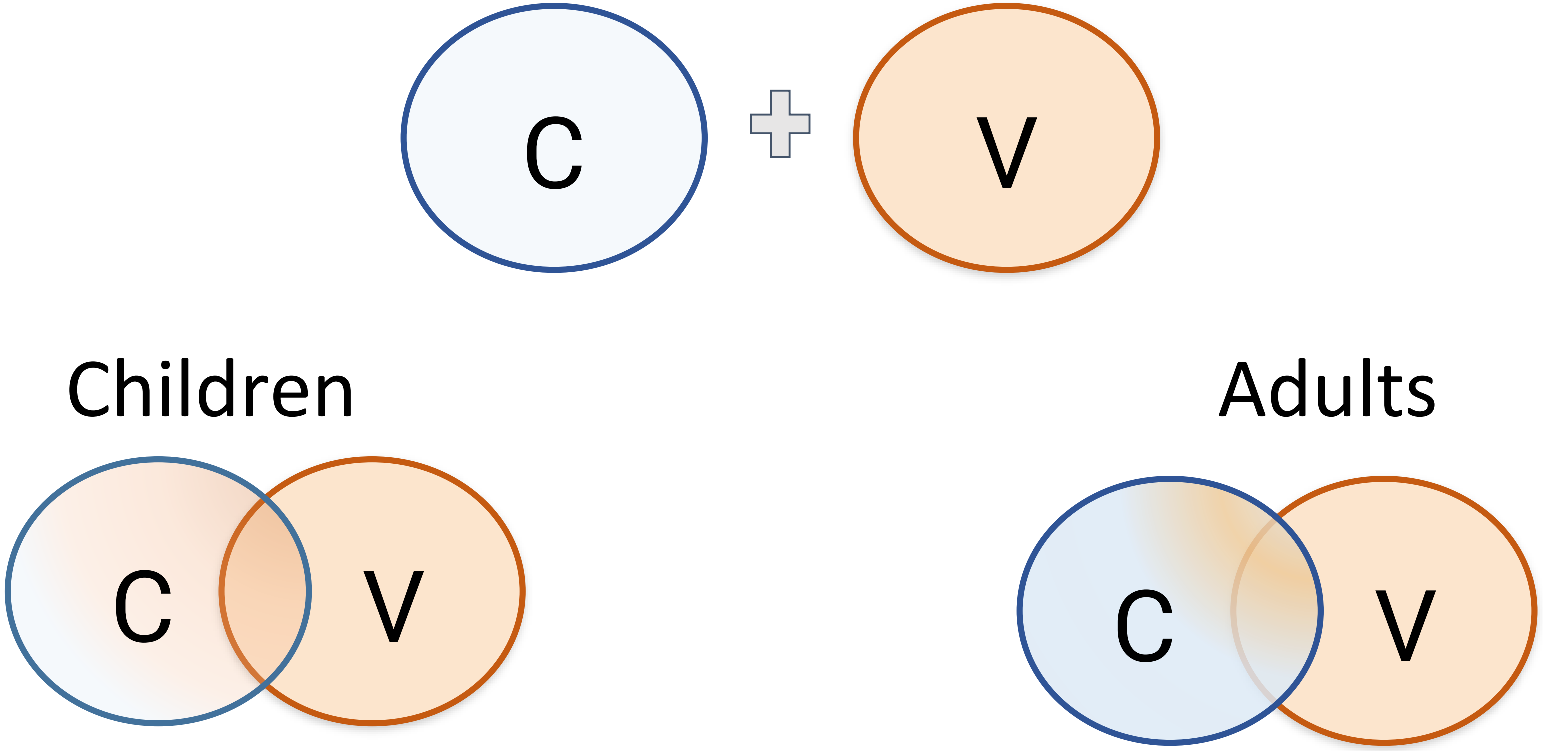
Hicks et al. 2015

Simulation experiment | Step 1



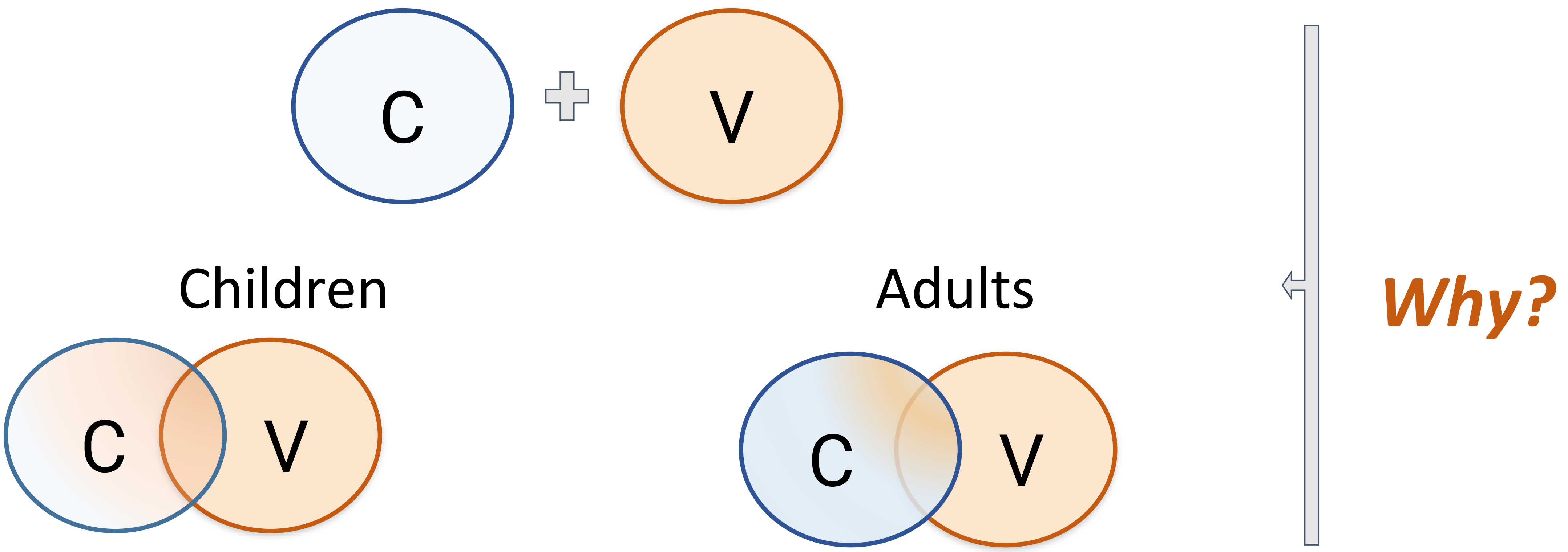
Nittrouer, Studdert-Kennedy, & Neely, 1996; Noiray et al. 2018

Simulation experiment | Step 1



Nittrouer, Studdert-Kennedy, & Neely, 1996; Noiray et al. 2018

Simulation experiment | Step 1



Nittrouer, Studdert-Kennedy, & Neely, 1996; Noiray et al. 2018

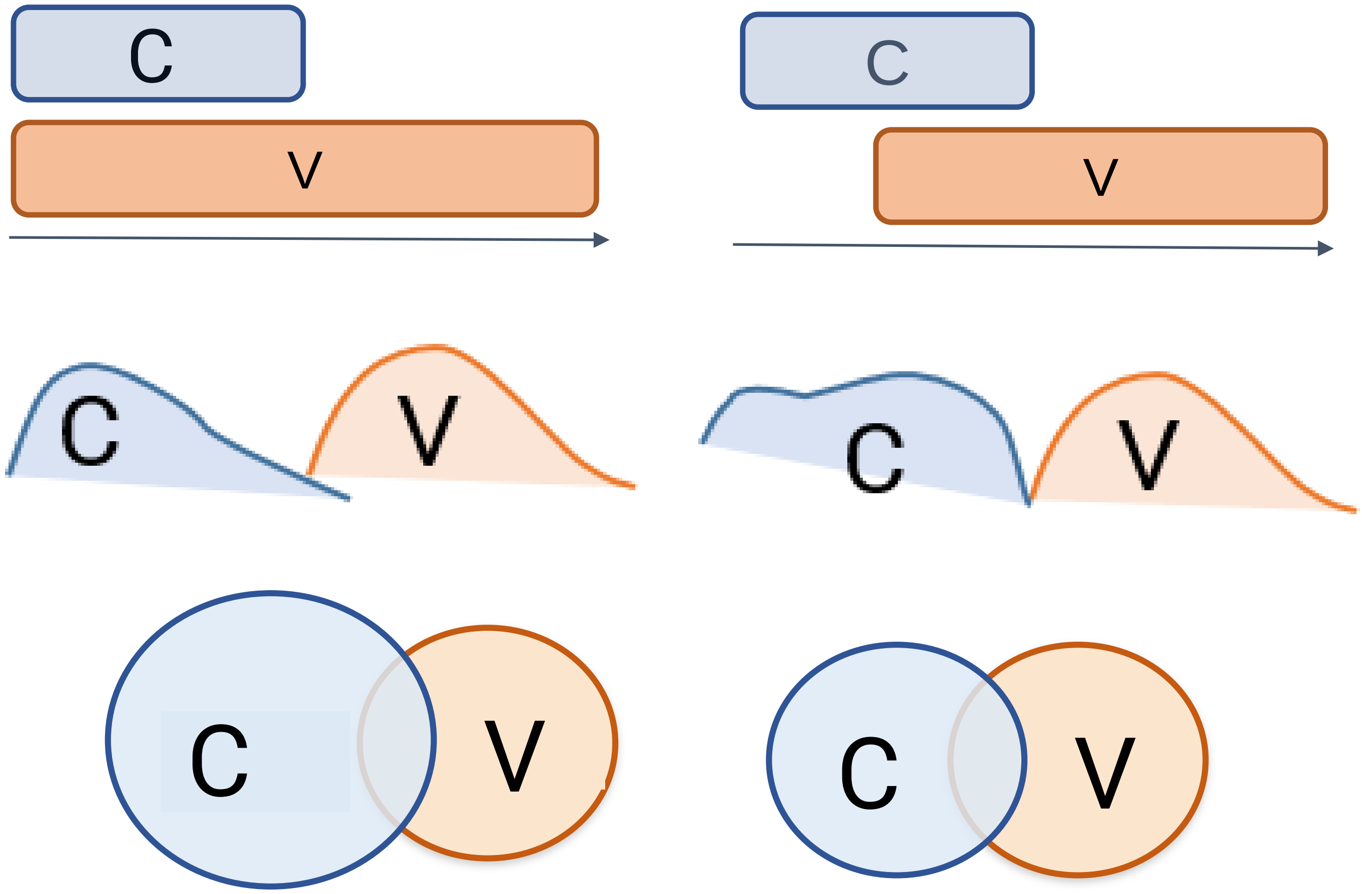
Simulation experiment | Step 1

Children > Adults

Timing

Articulatory strategies

Representations / variability



Step 1: Research Question

1. Are modeling and simulation necessary?

Yes, because there is no way to directly measure variability in representatioons

2. Is it possible to test your hypothesis with a model or simulation?

I think, yes

3. Is there an existing modeling and simulation framework that is capable of answering your research question? If not, do you have the expertise and resources to build one?

Yes, TADA

Step 1: Research Question

Observation: Children show more coarticulation in CV sequences than adults

Question: Does variability in articulator realization explain differences in coarticulation?

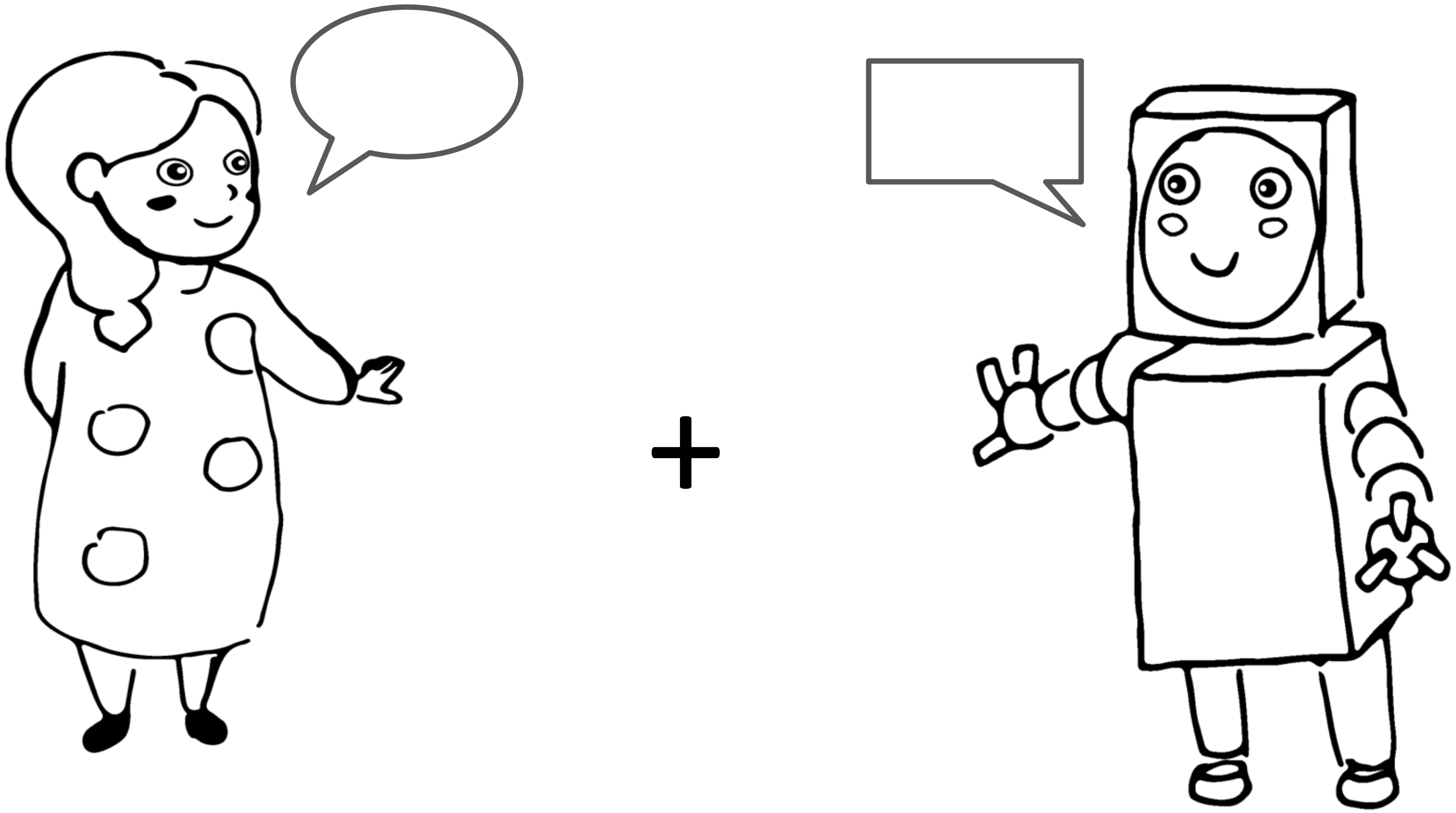
Hypothesis: Variability in representations is responsible for developmental differences in coarticulation

Method: Ultrasound data + articulatory modeling

Step 2: Make a plan

- Obtain experimental data for calibration → ultrasound dataset
- Manipulate blending value (determines coarticulation degree)
- To decide whether simulation is good enough, we are going to compare plots of tongue contours, exp vs sim
- Validation: use independent ultrasound dataset

Simulation experiment | Step 2



Step 3 Experimental data

- midsaggital tongue contours
- ultrasound imaging
- repetition task

Subjects: native speakers of German

3-5y.o.	7y.o.	Adults
n=48	n=14	n=11
25 females	11 females	6 females

Stimuli: pseudowords

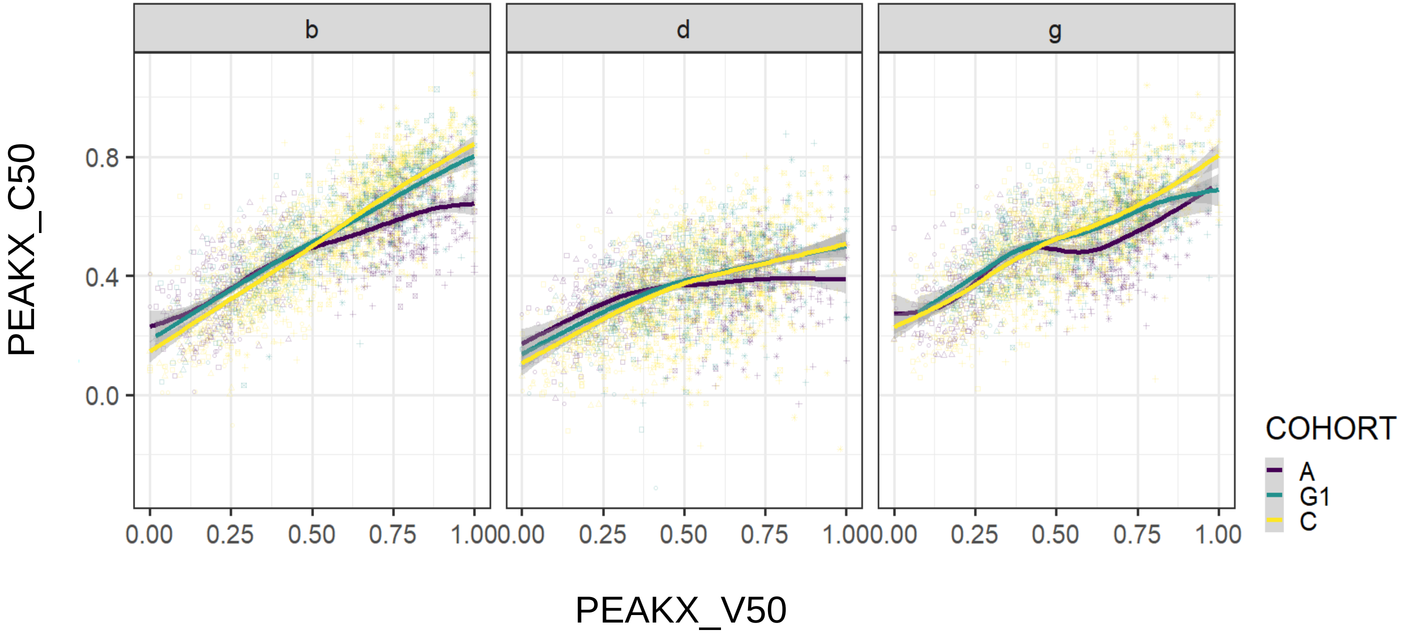
/i,e,y,a,o,u/

/b,d,g/

aɪnə $C_1 V C_2$ə

Noiray et al. 2019

Step 3 Experimental data



PEAKX – normalized horizontal position of the highest point on the tongue body
C50 - consonant midpoint; *V50* - vowel midpoint
A – adults, G1 – first-grade kids, C - preschoolers

- Coarticulation degree – here, the effect of tongue position during vowel production on the tongue position during consonant production
- Higher slope = higher coarticulation degree
- For all consonant,

CD in C > CD in G1 > CD in A

Step 4 Calibration

- finding the best match to experimental data.

What to manipulate?
Model parameter to manipulate – alpha value (blending) in TV.O files

How?
Depending on the model / question, can be done by
- Reverse engineering, or
- Based on specific values

For this demonstration, let us try several values for consonant blending.

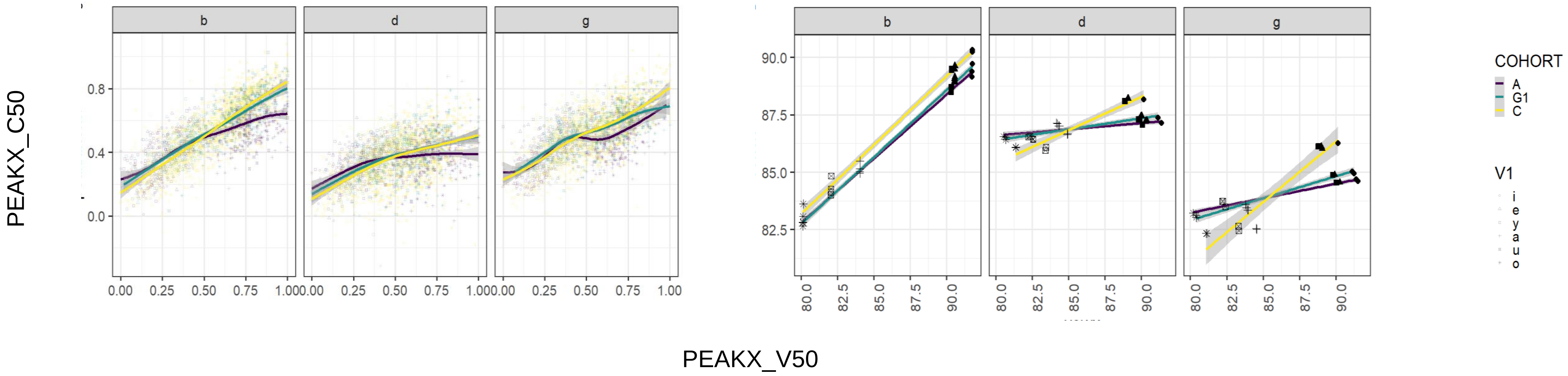
Step 4 Calibration

Script for this

- Adjust the script to run different values of blending
- The script will generate a table with all model articulator positions for all frames
- R plot comparing experimental data with simulated data

Step 4 Calibration

Experimental data **Simulated data**



- Normally one would make a quantitative comparison, but for now let us examine the plots
- Simulations show the same trend in general, however, too little difference for /b/ and too big of a difference for /g/

-> Keep calibrating

Step 5 Validation, or Is my model generalizable?

1. Compare simulations to independent data, such as
    1. Same experiment, different methods (articulatory data vs. acoustic data)
    2. Someone else's data
    3. Data from your previous experiments.

2. Compare simulations to other models and simulations

Step 6 Document and Share

Common structure for a simulation paper

3. Methods
    3.1. Experimental data: describe data used as reference for simulations
    3.2. Simulations
        3.2.1. Describe the model
        3.2.2. Describe the calibration
        3.2.3. Describe the validation
4. Results

Thank you!

Resourses

- Tutorial on task dynamics and articulatory phonology from Elliot Saltzman https://www.purdue.edu/tislr10/pdfs/Saltzman.pdf
- Course by Louis Goldstein at USC. Very informative https://sail.usc.edu/~lgoldste/ArtPhon/
- Dynamical System Simulations

https://demonstrations.wolfram.com/OrbitalDynamicsInFieldOfTwoPlanets/

Acknowledgements?

This work was supported by the DFG GZ: NO 1098/2-1 and 255676067.
We thank LOLA team for their assistance at various stages of the project.
Special thanks go to our adult and child participants (and their parents)!

https://demonstrations.wolfram.com/OrbitalDynamicsInFieldOfTwoPlanets/

References

- Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, *6*(2), 201–251. https://doi.org/10.1017/S0952675700001019
- Hicks, J. L., Uchida, T. K., Seth, A., Rajagopal, A., & Delp, S. L. (2015). Is my model good enough? Best practices for verification and validation of musculoskeletal models and simulations of movement. Journal of biomechanical engineering, 137(2), 020905. https://doi.org/10.1115/1.4029304
- Iskarous, K., Goldstein, L., Whalen, D., Tiede, M., & Rubin, P. (2003). CASY: The Haskins Configurable Articulatory Synthesizer. https://www.semanticscholar.org/paper/CASY%3A-The-Haskins-Configurable-Articulatory-Iskarous-Goldstein/b698c5024137f41e3852c787f588691ab5cc5a12
- Nam, H., Goldstein, L., Saltzman, E., & Byrd, D. (2004). TADA: An enhanced, portable Task Dynamics model in MATLAB. *The Journal of the Acoustical Society of America*, *115*(5), 2430. https://doi.org/10.1121/1.4781490
- Parrell, B., Lammert, A. C., Ciccarelli, G., & Quatieri, T. F. (2019). Current models of speech motor control: A control-theoretic overview of architectures and properties. The Journal of the Acoustical Society of America, 145(3), 1456-1481.
- Popescu (2019): Temporal organization of liquids in complex syllables
- Simmering VR, Triesch J, Deák GO, Spencer JP. To Model or Not to Model? A Dialogue on the Role of Computational Modeling in Developmental Science. Child Dev Perspect. 2010 Aug;4(2):152-158. doi: 10.1111/j.1750-8606.2010.00134.x.
- Tilsen, S. (2016). Selection and coordination: The articulatory basis for the emergence of phonological structure. *Journal of Phonetics*, *55*, 53-77.
- Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. Ecological Psychology, 1(4), 333–382. https://doi.org/10.1207/s15326969eco0104_2

Further reading

- Parrell, B. (2011). Articulatory Phonology and the Task Dynamics Model. *Studies in Hispanic and Lusophone Linguistics*, *4*(2), 531-542. https://doi.org/10.1515/shll-2011-1112
- Iskarous, K. (2017). The relation between the continuous and the discrete: A note on the first principles of speech dynamics. *Journal of Phonetics*, *64*, 8-20.