Two Tools for Learning Gestural Parameters and Constraint Based Grammars

Computational Approaches to Phonology
2022 Annual Meeting on Phonology
October 21, 2022

Caitlin Smith
University of North Carolina at Chapel Hill

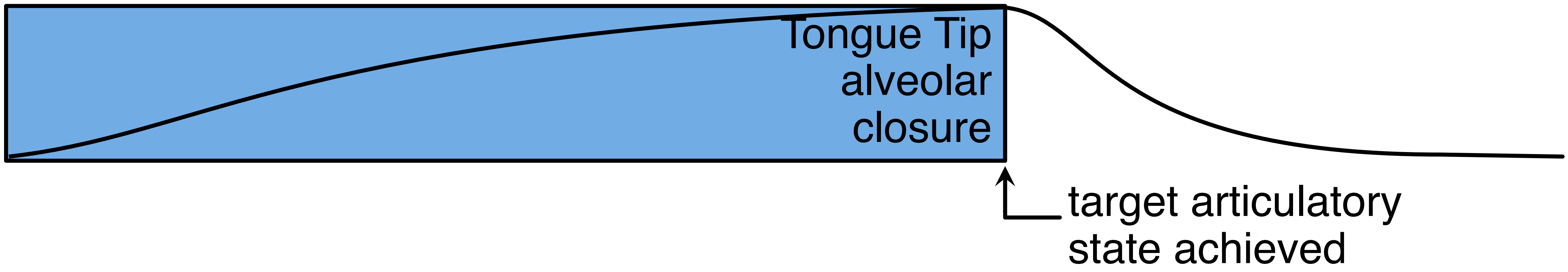
Charlie O'Hara
University of Michigan

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

UNIVERSITY OF MICHIGAN

Gestures in Articulatory Phonology
(Browman & Goldstein 1986, 1989 et seq.)

- Gestures: dynamically-defined, goal-based units of phonological representation in Articulatory Phonology

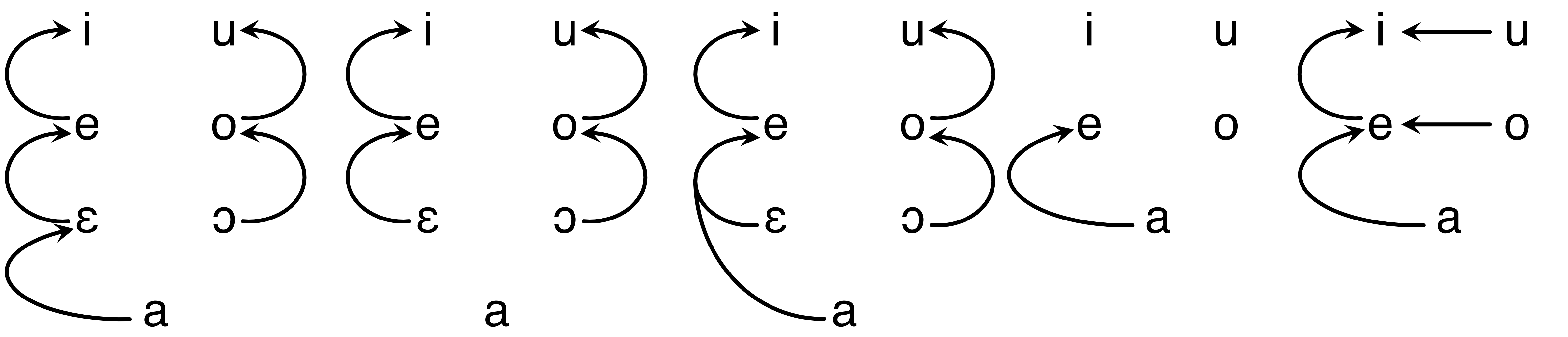


- Ample work on computational modeling of speech production with gestural representations, e.g. TADA (Task Dynamic Application; Nam et al. 2004)
- Less work on computational modeling of learning of gestural parameters from observed articulations

Chain-Shifting Height Harmony

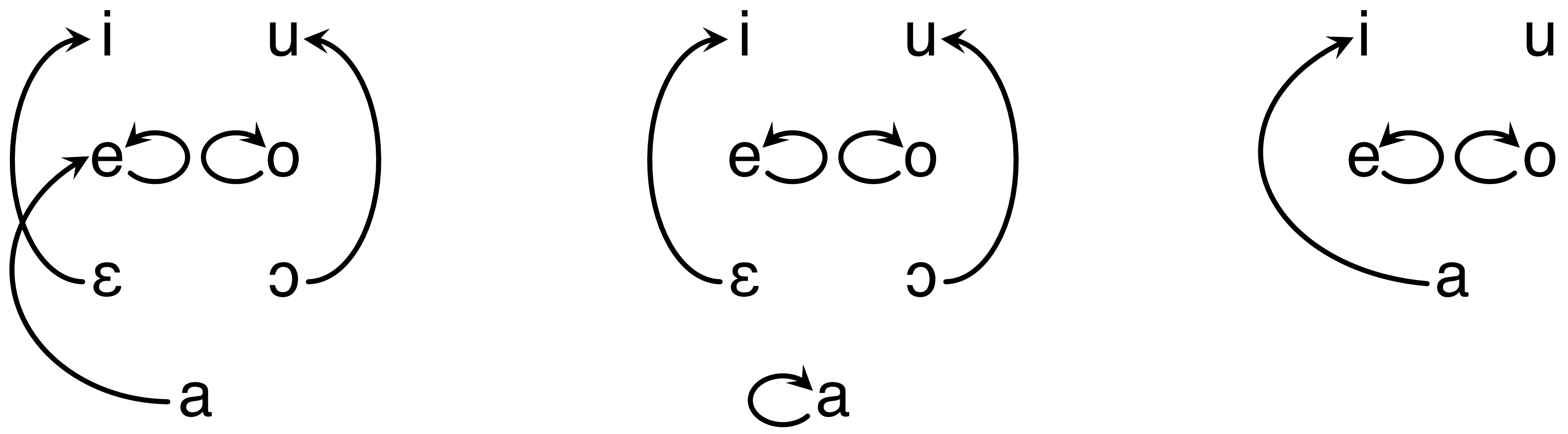
Chain-shifting vowel raising patterns in which vowels raise single step along height scale are well attested:

Nzebi	Servigliano Italian	Basaa	Basque	Fyem
Guthrie (1968)	(Nibert 1998)	(Schmidt 1996)	(Hualde 1988)	(Nettle 1998)



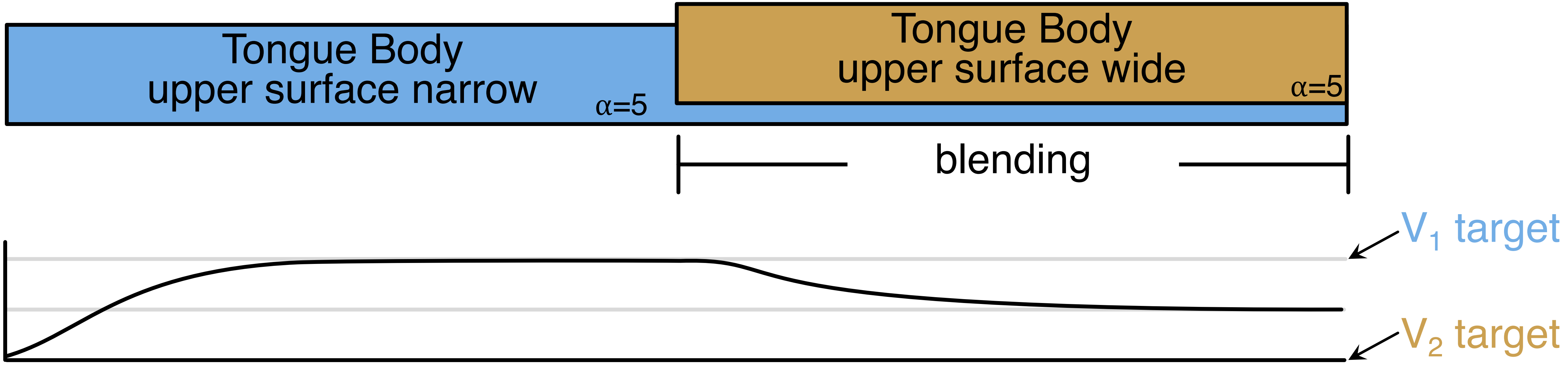
Unattested Saltatory Height Harmony

Two-step vowel raising patterns that ‘skip over’ a step in the height scale (i.e., saltation) are unattested (Parkinson 1996):



Gestural Overlap, Blending, and Strength

- Gestural Harmony Model (Smith 2016, 2018, et seq.): harmony is extension of trigger gesture to **overlap** surrounding gestures
- Partial height harmony is result of **blending** between vowel gestures with different target articulatory states (heights)



Gestural Overlap, Blending, and Strength

- Conflicting gestural targets resolved by blending targets of overlapped gestures according to Task Dynamic Model of speech production (Saltzman & Munhall 1989, Fowler & Saltzman 1993)

$$\frac{\text{Target}_1 \times \text{Strength}_1 + \text{Target}_2 \times \text{Strength}_2}{\text{Strength}_1 + \text{Strength}_2} = \text{Blended Target}$$

- Different height harmony patterns results from different gestural parameter settings for vowels' target heights and blending strengths

Gestural Gradual Learning Algorithm (GGLA)

(Smith & O’Hara 2021)

- Learns gestural parameters necessary to correctly produce target partial height harmony pattern (chain shift or saltation) when trigger/undergoer overlap occurs
- Agent-based: learner agent trying to learn harmony pattern produced by teacher agent
- Online and error-driven: For each incorrect vowel production, perform parameter updates to any vowel gestures involved in that production

Error-Driven Online Learning

- During a learning trial, randomly generated two-syllable input is produced according to teacher's and learner's individual gestural parameter settings
- Check learner and teacher productions for a match (i.e. both learner vowels produced within specified window around both teacher vowels)
- In event of production error, perform necessary updates to learner's constraint weights and/or gestural parameter settings (constriction degree, blending strength)

Gestural Gradual Learning Algorithm (GGLA)

(Smith & O'Hara 2021)

- Constriction degree update for blended and unblended vowel productions:

sgn(teacher prod. – learner prod.) × LR

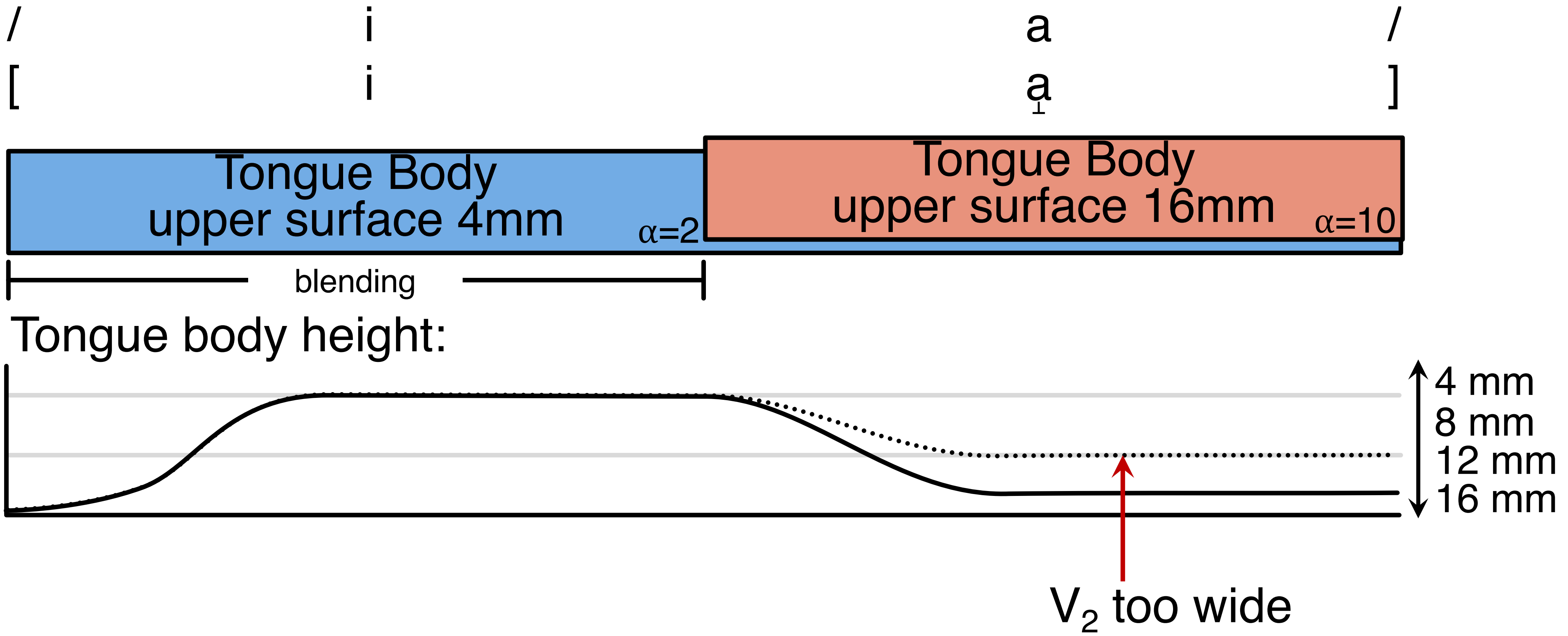
- Blending strength updates for blended vowel gestures V_1 and V_2:

V_1: sgn(learner V_1 CD – learner V_2 CD) × sgn(teacher prod. – learner prod.) × LR

V_2: sgn(learner V_2 CD – learner V_1 CD) × sgn(teacher prod. – learner prod.) × LR

CD = constriction degree, LR = learning rate

Sample Training Iteration



/i/ updates: 3.9mm↓
2.1α↑

/a/ updates: 15.9mm↓
9.9α↓

GGLA on Google Colab

Visit the Google Colab for the GGLA:

tinyurl.com/ampggla

Extending the GGLA

- GGLA-based learning simulations assume:
  - Learner already knows which vowels are harmony triggers
  - All vowels are undergoers of harmony
- GGLA-based learning simulations do not assume:
  - Any kind of phonological grammar
  - Any kind of input-output mapping

Grammar + Gesture Gradual Learning Algorithm (GGGLA) (Smith 2022)

- Compares articulatory productions of gestural scores from teacher and learner agents during training
- Learns phonological grammar necessary to correctly cast vowels as triggers, undergoers, and blockers of harmony
- Learns gestural parameters necessary to correctly produce a given height harmony pattern (chain shift or saltation) when trigger/undergoer overlap occurs via GGLA

Constraint Weight Learning in Maximum Entropy Harmonic Grammar

- Constraint-based grammars implemented in Maximum Entropy Harmonic Grammar (MaxEnt; Goldwater & Johnson 2003; Jäger 2007): weighted constraints and variable outputs

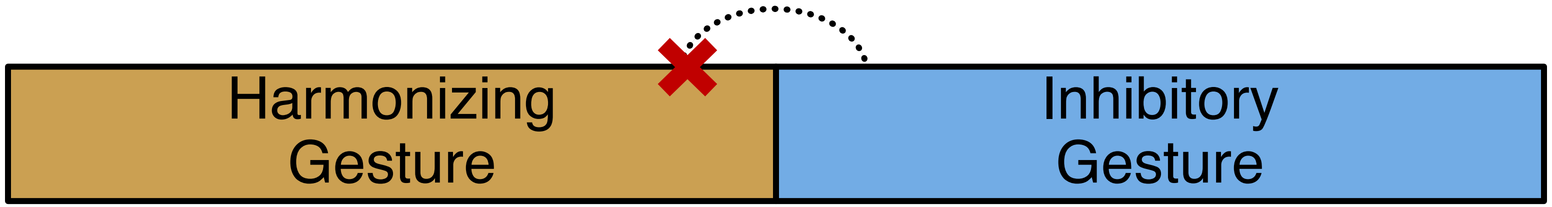
- Learning of constraint weights based on Gradual Learning Algorithm for Harmonic Grammar (GLA; Boersma & Pater 2016), based on earlier Perceptron Algorithm (Rosenblatt 1958)

Error-Driven Online Learning (Again)

- During a learning trial, randomly generated two-syllable input is evaluated by teacher's and learner's individual phonological grammars and winning output candidate gestural score is produced according to their individual gestural parameter settings
- Check learner and teacher productions for a match (i.e. both learner vowels produced within specified window around both teacher vowels)
- In event of error, perform necessary updates to learner's constraint weights and/or gestural parameter settings (constriction degree, blending strength)

Blocking as Gestural Inhibition

- Blocking: inhibition (deactivation) of harmonizing gesture by incompatible gesture, preventing gestural overlap (Smith 2016, 2018)

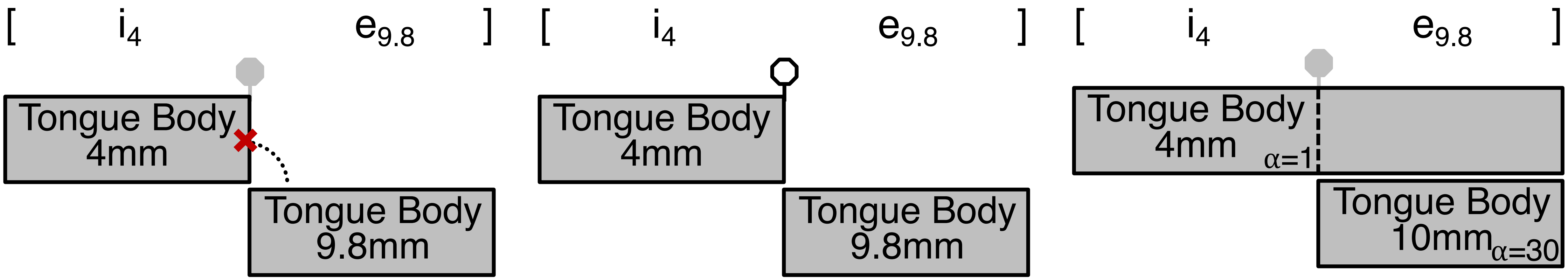


- Inhibited gesture cannot reactivate itself, so harmony ceases
- Combined effects of blending and blocking result in productions whose gestural underpinnings are ambiguous

Articulatory Productions and Hidden Structure

- To perform constraint weight updates, learner must compare violations incurred by its and teacher's chosen winner candidates

- Teacher's output candidate gestural score is often unknown to learner based solely on teacher's production:



Hidden Structure Learning

- Robust Interpretive Parsing (RIP; Tesar & Smolensky 1998, Boersma 2003): determine teacher candidate's hidden structure (e.g. metrical structure) from its surface form

- Robust Interpretive Production Parsing (RIPP): determining teacher candidate's hidden phonological surface form (i.e. gestural score) from its articulatory production

GGGLA on Google Colab

Visit the Google Colab for the GGGLA:

https://tinyurl.com/ampgggla

Conclusions

- Results of learning simulations:
  - With GGLA, extreme strengths necessary to generate saltatory height harmony are substantially slower/harder to learn
  - With GGGLA, extreme strengths and/or grammars favoring blocking (no gestural overlap) necessary to generate saltatory height harmony are substantially slower/harder to learn

- Learning simulations correctly indicate a typological bias favoring attested chain-shifting harmony and against saltatory harmony

Works In Progress / Coming Soon

- Learning of additional gestural parameters (e.g. constriction location, persistence)
- Learning of additional phonological phenomena (e.g. consonant lenition)
- Support for multi-gestural segments to learn settings of multiple parameters for multiple phenomena concurrently
- Noisy productions of output gestural representations
- Running GGLA and GGGLA with graphical user interfaces