From Gestural Landmarks to Analysis-by-Synthesis: Tone-driven Gestural Timing in Tibetan

Christopher Geissler

Heinrich-Heine-University Düsseldorf, Germany

{christopher.geissler}@hhu.de

Introduction

- Aim 1: Use simulation to test the hypothesis that **tone gestures** affect timing of oral articulatory gestures
- Aim 2: Provide proof-of-concept for testing **temporal coordination** with articulatory simulation in TADA

Simulating articulation in the Task Dynamics Application

- TADA: "TAsk Dynamics Application" [5]
- Words + gestural dictionary + syllable structure → gestural score
  gestural score → articulatory trajectories → acoustics
- Highly-parameterized model can be edited at any stage

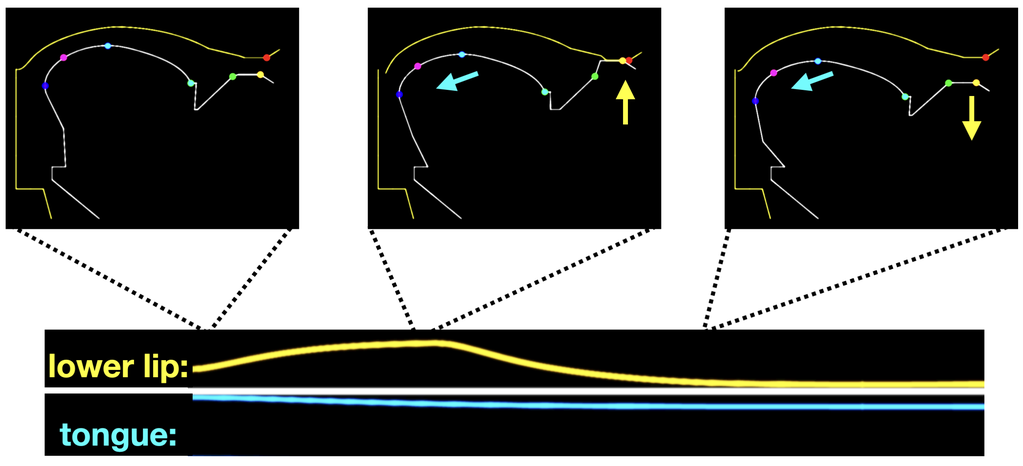
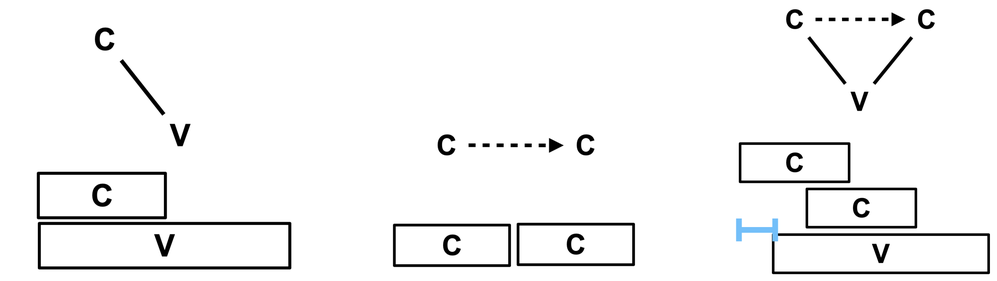


Figure: Simulated articulatory trajectories of lower lip and tongue dorsum in [ma], with in-phase coupling

Coupling modes and tone gestures

- Representations in Articulatory Phonology
  - Gestures: constriction degree, constriction location, duration
  - Coupled Oscillator Model of Syllable Structure
  - Task Dynamics: how implement gestures in movement
- Gestures self-organize according to coupling modes: **in-phase** / synchronous, **anti-phase** / sequential. Combining gestures in **competitive coupling** results in partial overlap; explains "C-Center" effects [5].



(a) In-phase coupling C-V; T-V?

(b) Anti-phase coupling: C-C, V-V, V-C

(c) Competitive coupling: CCV; CTV?

- Tone gestures: in lexical tone languages, CV syllables have timing reminiscent of cluster. Explained by treating tone as a gesture with consonant-like timing [1], [4], [3]
- **Prediction:** CV syllable with tone will have C-V timing like that of CCV syllable, because of competitive coupling between C and tone gestures

EMA Data

- Electromagnetic articulography data of Tibetan speaker living in US (from [2])
- Current: syllable [ma] with high tone (language also has a rising tone)

Acknowledgements

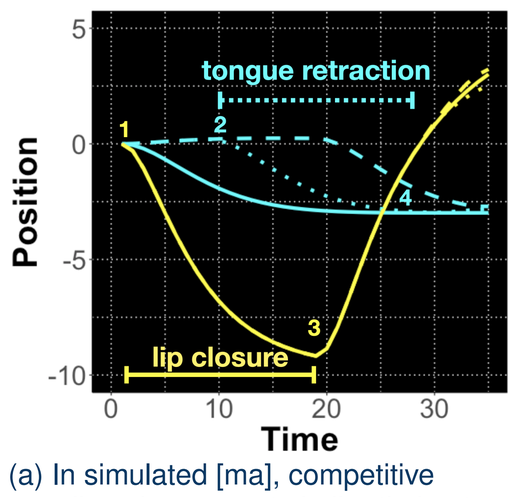
Thanks to Kevin Tang, Jason Shaw, and Mark Tiede for advice and encouragement; Muye Zhang and Jason Shaw for data collection help; and the Tibetan speakers for their time and voices. Funding for experiments from the Phonetics Lab at Yale University. Thanks also to practice audiences at the Speech, Language, and Modeling (SLaM) Lab at HHU.

Simulation procedure

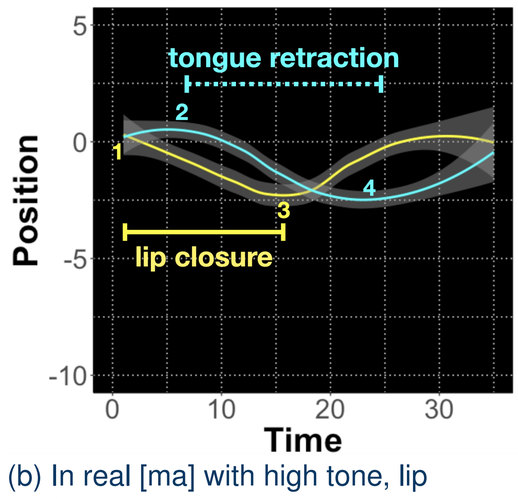
- Generate [ma] in TADA, then manually edit the timing of gestures reflecting presence and absence of hypothesized tone gesture
- Create three versions of [ma]: in-phase C-V (non-tonal); anti-phase C-V; and competitive-coupling C-V timing as predicted by tone gesture hypothesis

Results: looks like competitive coupling

- Subtracted position of each datapoint from value at first timestamp
- Tongue dorsum retraction begins in the middle of the lip closure, and lip closure ends around the middle of tongue dorsum retraction. This most closely resembles **competitive coupling**, and supports the tone gesture hypothesis.



(a) In simulated [ma], competitive coupling shows most similar timing to real data; solid = in-phase, dotted = competitive coupling, dashed = anti-phase.



(b) In real [ma] with high tone, lip gesture is smaller, gestures are slightly shorter, and movement toward targets in the next word is visible. Timing resembles competitive coupling

Discussion

- Relative timing of gestures is most consistent with competitive coupling
- TADA simulations are close enough to observed data to be instructive
- Next steps: do more! More languages, more syllable types, more complex tones.
- Long term: interested in automating the editing of gestural score files, with the goal of making gestural annotations for EMA data, building on [6]

References

[1] Man Gao. "Tonal alignment in Mandarin Chinese: An articulatory phonology account". In: *Unpublished Doctoral Dissertation (Linguistics), Yale University, CT* (2008).

[2] Christopher Geissler. "Temporal articulatory stability, phonological variation, and lexical contrast preservation in diaspora Tibetan". PhD thesis. Yale University, 2021. URL: https://elischolar.library.yale.edu/gsas_dissertations/52.

[3] Christopher Geissler et al. "Eccentric CV timing across speakers of diaspora Tibetan with and without lexical tone contrasts". In: *Proceedings of the 12th International Seminar on Speech Production.* 2021.

[4] Robin Karlin. "The articulatory TBU: gestural coordination of tone in Thai". In: *Cornell Working Papers in Linguistics*. 2014.

[5] Hosung Nam and Elliot Saltzman. "A competitive, coupled oscillator model of syllable structure". In: *Proceedings of the 15th International Congress of the Phonetic Sciences*. 2003.

[6] Hosung Nam et al. "A Procedure for Estimating Gestural Scores from Natural Speech". en. In: *Interspeech 2010* (2010), p. 4.