The phonology of articulatory coordination

Christopher Geissler
Heinrich Heine University Düsseldorf

April 10, 2023
Carleton College Linguistics

Roadmap

- Introduction
- Articulatory Coordination
- Studies
  - Tibetan tone
  - English diphthongs
- Conclusion

Roadmap

- **Introduction**
- Articulatory Coordination
- Studies
  - Tibetan tone
  - English diphthongs
- Conclusion

Things this talk isn’t about

but I’m also interested in

- Swarthmore, ELA, Yale, HHU
- Gender bias in linguistics examples
- L1 transfer (German-to-English)
- Probabilistic reduction
- Scholarship of teaching

Where I am from?

- Contrast *cot* vs. *caught* and *Don* vs. *dawn*

  [kʰɑt] [kʰɔt] [d̥ɑn] [d̥ɔn]

- Contrast *Mary* vs. *marry* vs. *merry* vs. *Murray*

  [ˈmɛ͡ɪ.ɹi] [ˈmæ.ɹi] [ˈmɛ.ɹi] [ˈmʌ.ɹi]

- Different vowels in *write* vs. *ride*, but not *house* (n) vs. *house* (v)

  [ɹʌ͡ɪt] [ɹa͡ɪd] [ha͡ʊs] [ha͡ʊz]

- The night before Halloween is “Mischief Night”
- I can either “wait in line” or “wait on line”

Some phonology

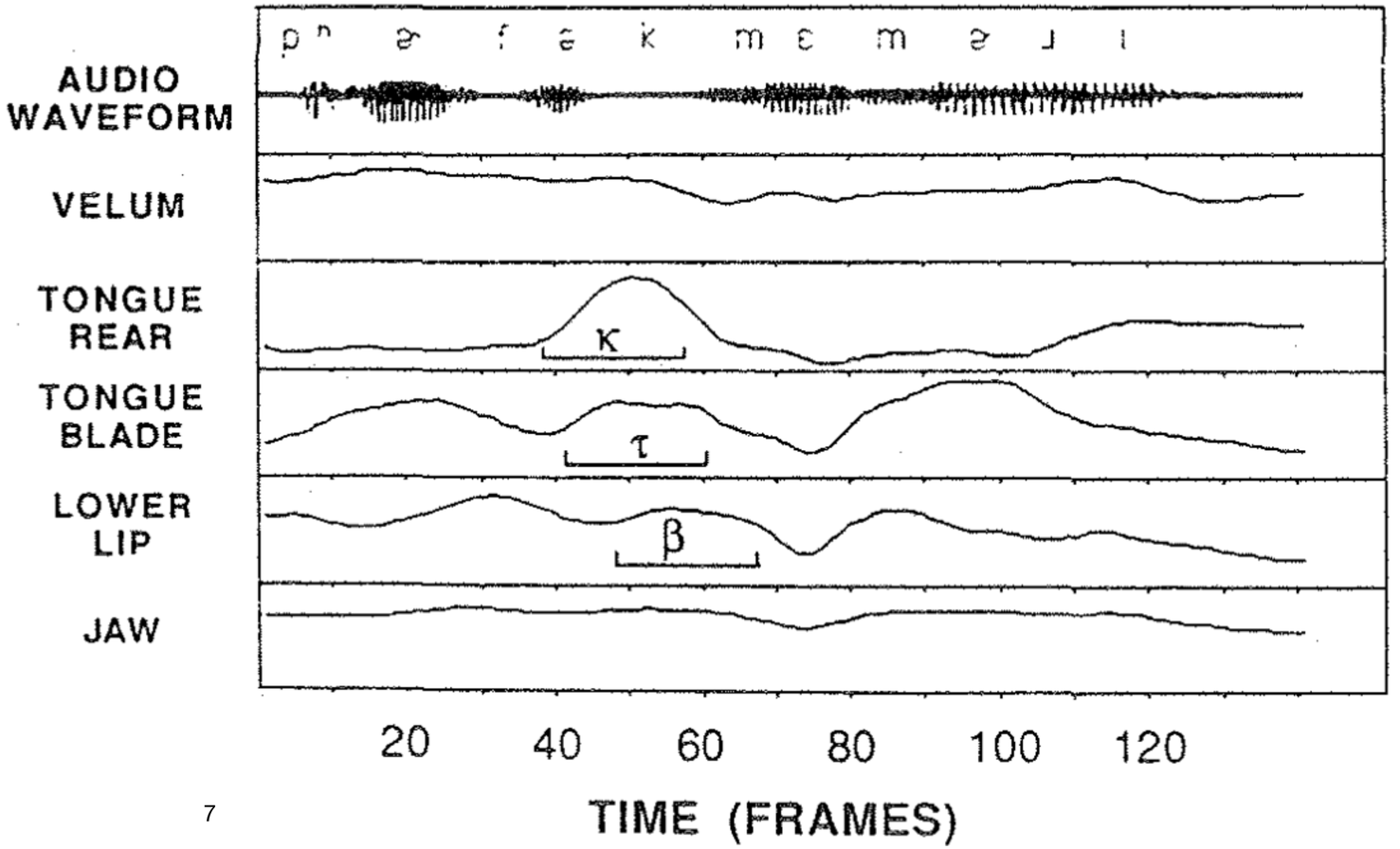
- *write* [ɹʌ͡ɪt] → *writer* [ɹʌ͡ɪ.ɾɻ]; *ride* [ɹa͡ɪd] → *rider* [ɹa͡ɪ.ɾɻ]
  - Raising: /a͡ɪ/ → [ʌ͡ɪ] / _[-voice]
  - Flapping: /t/ → [ɾ] / _[+syllabic,-stress]
- Rule ordering:

	writer	***rider***
UR	/ɹa͡ɪ.tɻ/	/ɹa͡ɪ.dɻ/
Raising	|ɹʌ͡ɪ.tɻ|	|ɹa͡ɪ.dɻ|
Flapping	|ɹʌ͡ɪ.ɾɻ|	|ɹa͡ɪ.ɾɻ|
SR	✅ [ɹʌ͡ɪ.ɾɻ]	✅[ɹa͡ɪ.ɾɻ]

	writer	***rider***
UR	/ɹa͡ɪ.tɻ/	/ɹa͡ɪ.dɻ/
Flapping	|ɹa͡ɪ.ɾɻ|	|ɹa͡ɪ.ɾɻ|
Raising	|ɹa͡ɪ.ɾɻ|	|ɹa͡ɪ.ɾɻ|
SR	❌ [ɹa͡ɪ.ɾɻ]	✅ [ɹa͡ɪ.ɾɻ]

…but that doesn't always work

- "t/d deletion": e.g. *perfec(t) memory*
- Rule deleting the /t/?
- but what about in articulation?



Midsagittal sections

(Browman & Goldstein 1988)

Simultaneous units

- Discrete, linearly-ordered units are not sufficient
  - Other simultaneous productions:
    - pitch, loudness: *Amelia?* vs. *Amelia.* vs. *AMELIA!!!!!*
    - simultaneity in sign languages
    - complex segments: /t͡ʃ/ in *ch*ur*ch*, [a͡ɪ] in *fi*ve

Timing as phonology

- Temporal coordination is necessary in phonology
- Evidence from patterns of variation in…
  - …tone:
    - Tibetan: [má] ‘injury’ vs. [mà] ‘mother’
  - …diphthongs:
    - English: [a͡ɪ] in *five*

Roadmap

- Introduction
- **Articulatory Coordination**
- Studies
  - Tibetan tone
  - English diphthongs
- Conclusion

CV coarticulation

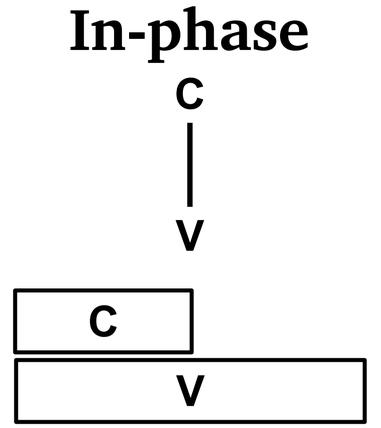
- CV syllable, like *tea* /ti/
  - [t] & [i] start simultaneously
  - [i] has longer duration
- VC syllable, like eat /it/
  - timing is sequential

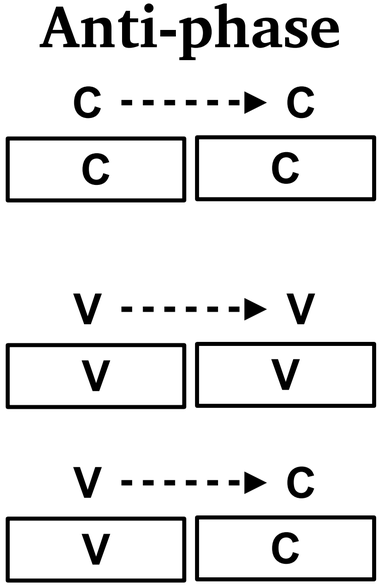
/ti/ ‘tea’	
LIPS	
TONGUE TIP	alveolar closure
TONGUE BODY	palatal narrow

/it/ ‘eat’	
LIPS	
TONGUE TIP	alveolar closure
TONGUE BODY	palatal narrow

CV coarticulation

- Two stable timing patterns:
  - In-phase = synchronous
  - Anti-phase = sequential
- Bimanual tapping example

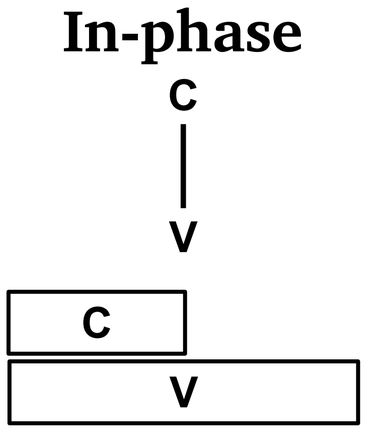


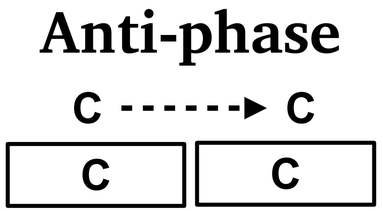


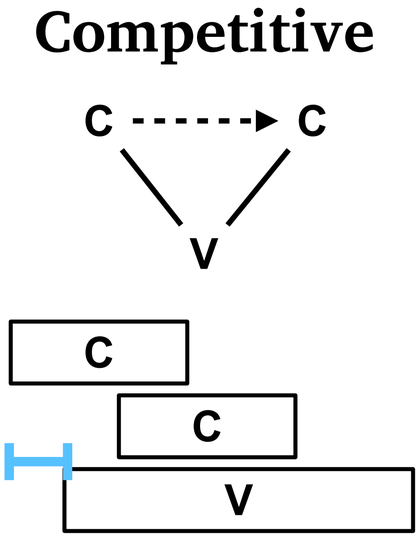
What about clusters?

- Partial overlap (“C-center”)
  - Competitive coupling

/spa/ ‘spa’	
LIPS	labial closure
TONGUE TIP	alveolar critical
TONGUE BODY	pharyngeal wide

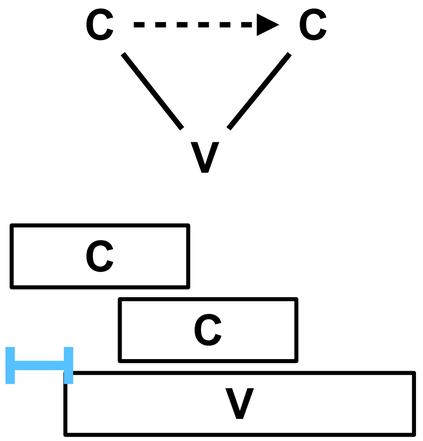




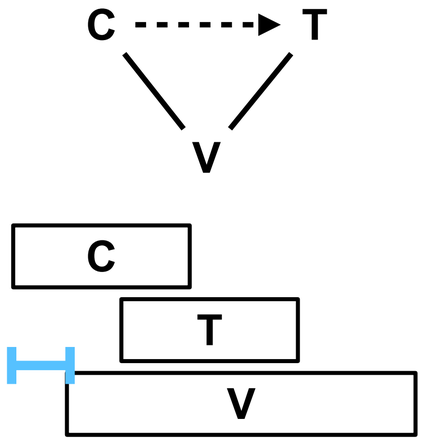


What about tone?

- phonologists argue about unit of tone:
  - on vowels? on syllables? … or its own gesture?



/spa/ ‘spa’	
LIPS	labial closure
TONGUE TIP	alveolar critical
TONGUE BODY	pharyngeal wide

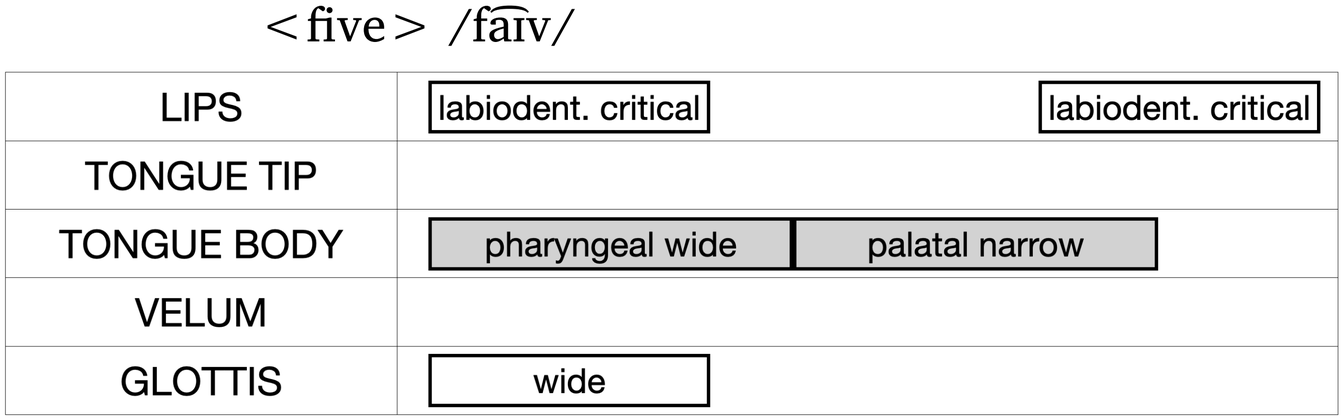


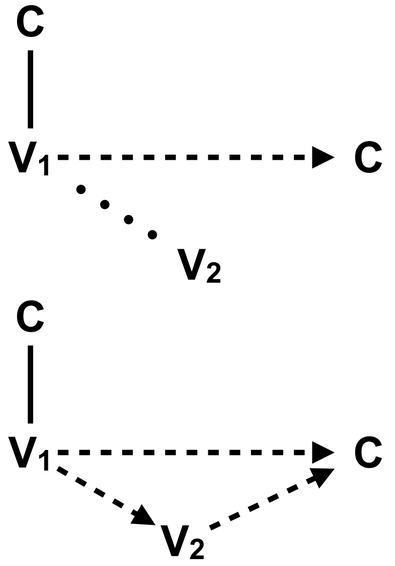
/pá/	
LIPS	labial closure
TONGUE TIP	
TONGUE BODY	pharyngeal wide
pitch (?)	high

(Gao 2008, Niemann et al. 2011, Karlin 2014)

What about diphthongs?

- Can approximately describe with in-phase/anti-phase
- How do diphthongs change when they get shorter?





(Goldstein et al. 2000)

Roadmap

- Introduction
- Articulatory Coordination
- **Studies**
  - **Tibetan tone**
  - English diphthongs
- Conclusion

Tibetan

- Dialects vary dramatically
  - tonal, non-tonal
  - different consonants, syllable structures
- Diaspora: mixed variety
  - Fieldwork: 73 interviews in Kathmandu (19 raised there)

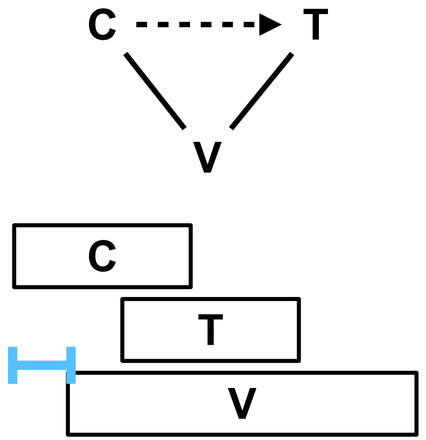
(Geissler 2019, 2021ch3)

  - Articulatory experiment: 6 participants in US

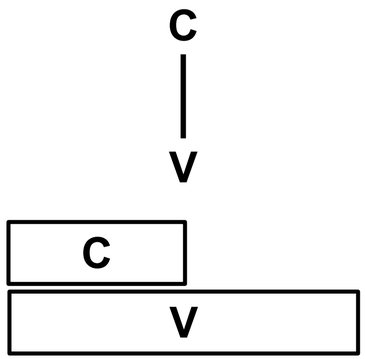
(Geissler et al. 2021, 2021ch4)

Predictions

- If there is a tone gesture in a syllable:
  - C-V timing like in clusters:
    C-V lag positive, ~50ms



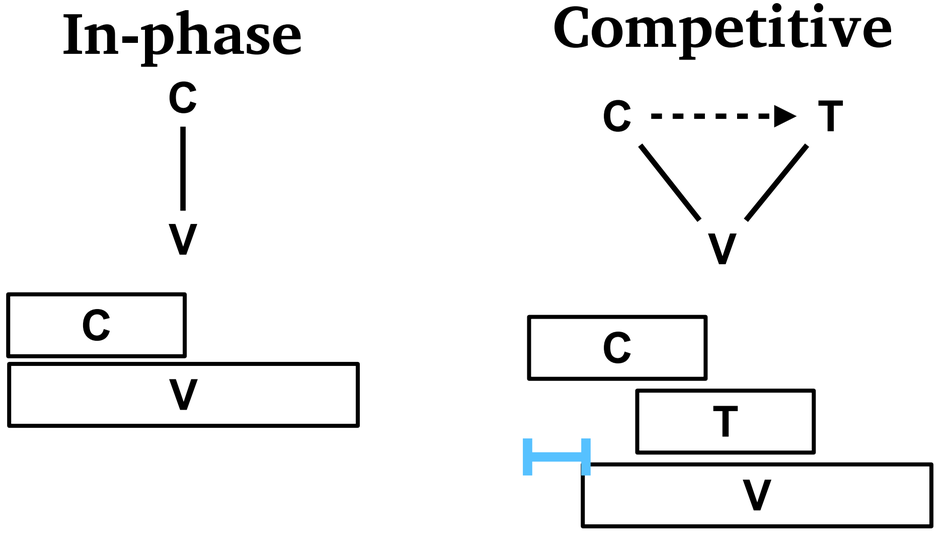
- If there is no tone in that syllable:
  - Simultaneous C & V:
    C-V lag ~0ms

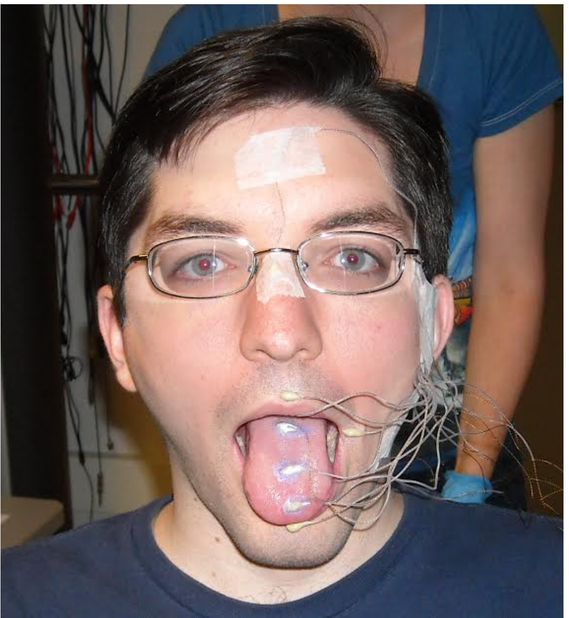


EMA study

articulatory trajectories

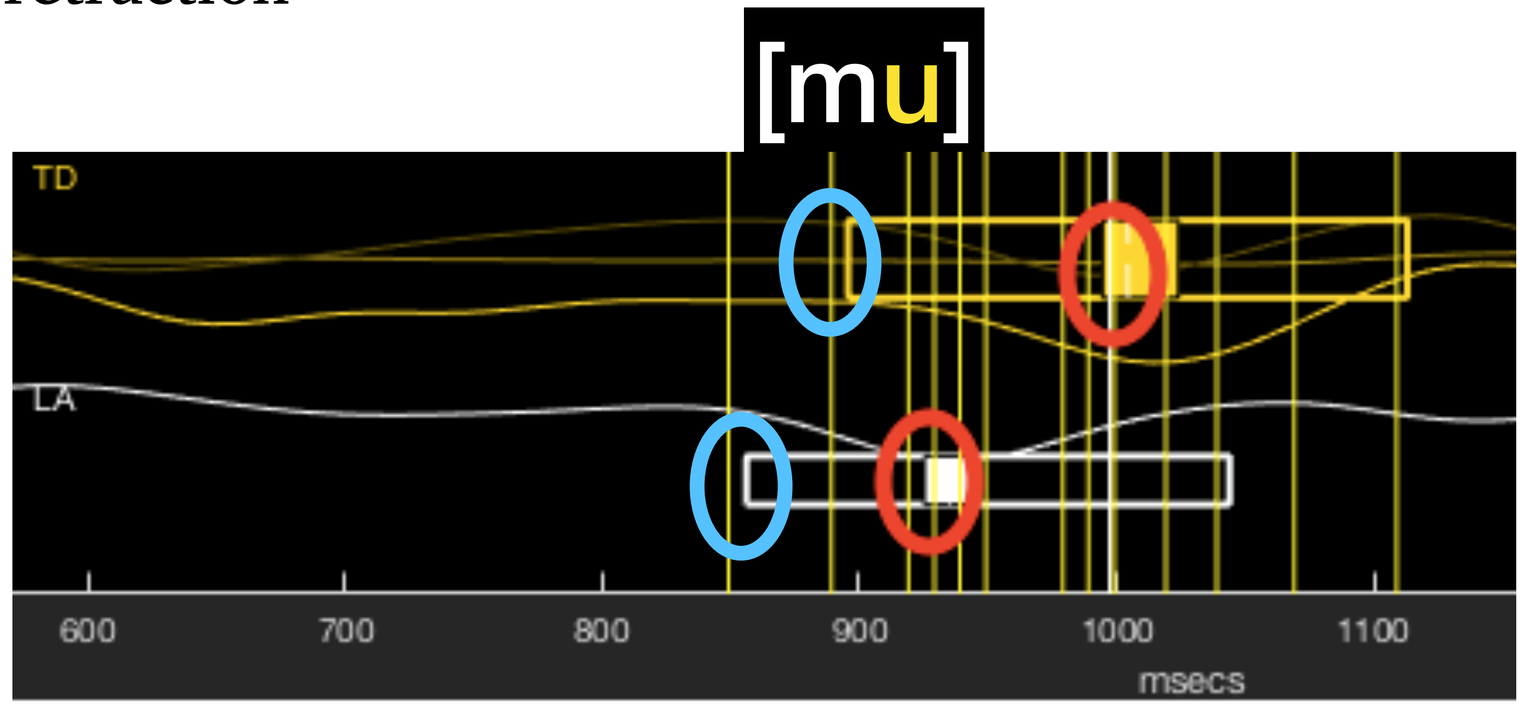
- Tracks movement of sensors over time
- [p p^h m]: distance between lip sensors
- [i]→[u o a]: tongue dorsum retraction
- C-V lag as diagnostic of tone





Tongue Dorsum: front ↓ back

Lip Aperture: open ↓ closed

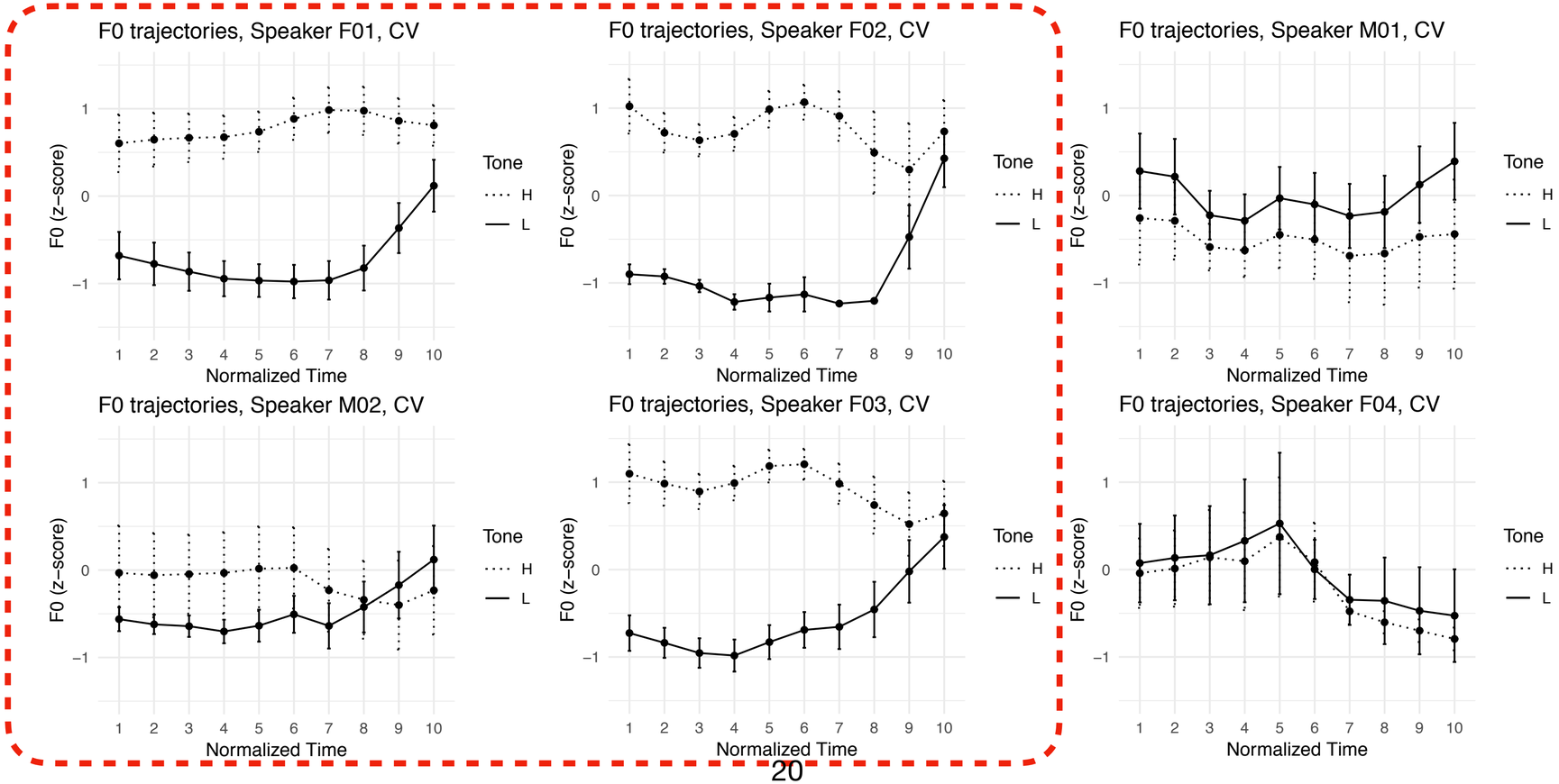


(Data: Zhang, Geissler, & Shaw 2019)
(Mview software: Tiede 2005)

Do the participants produce tone?

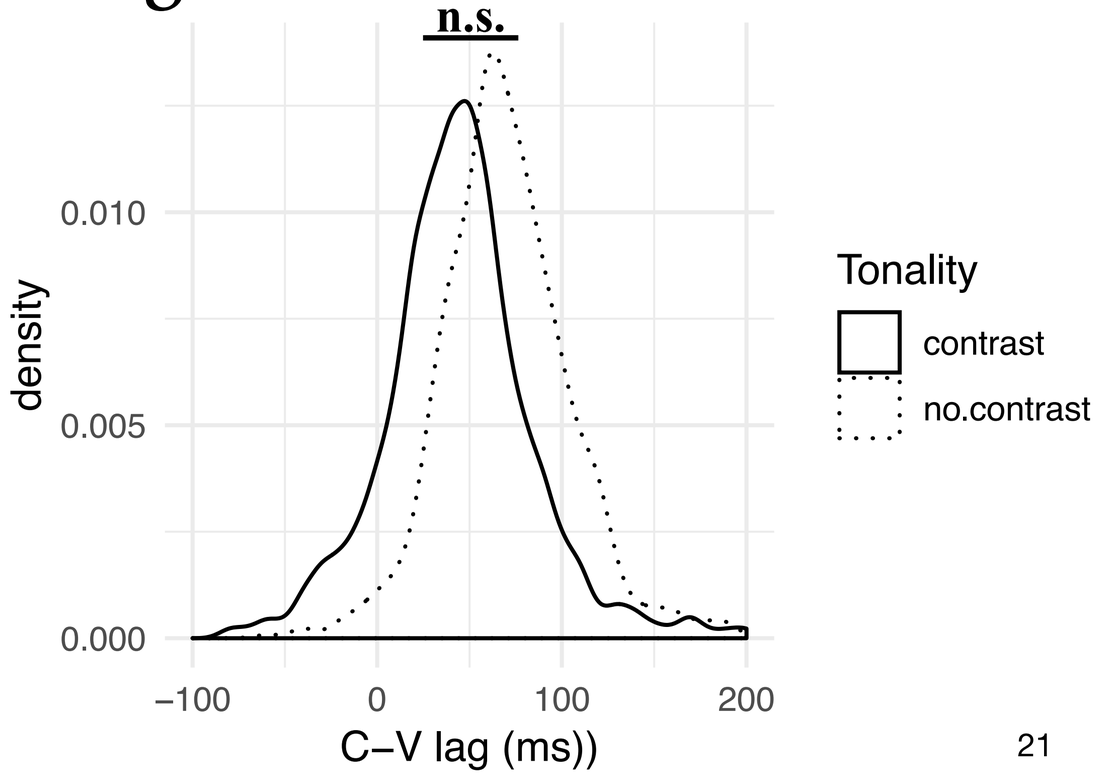
Evidence: pitch tracks

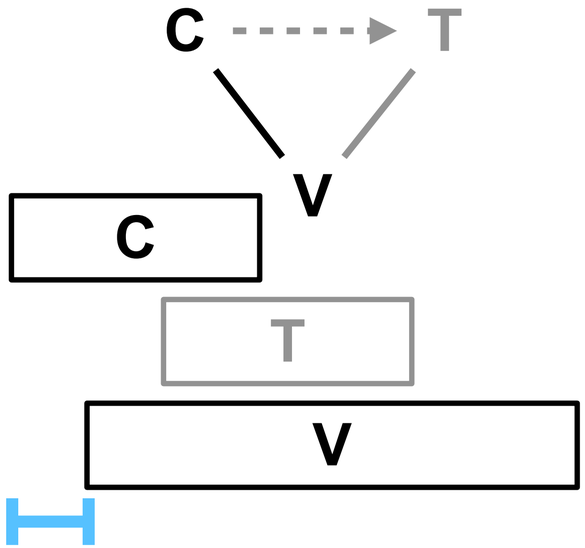
- 4 speakers produce a tone contrast, two do not (images: /mV/)



Results: C-V lag

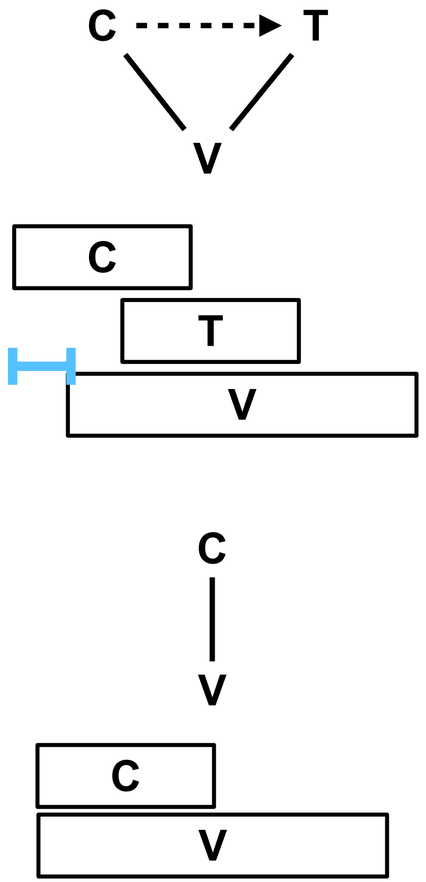
- There is a positive C-V lag… for speakers with ***and*** without the tone contrast
- No significant difference between the tones





Digging deeper

- C-V timing similar to a tonal syllable
- For non-tonal speakers… phonological “tone” should be persisting somehow
  - Where else might we find effects of “tone” 👻?



Two onset consonant systems

Both tonal *and* non-tonal (merged) speakers found with each

- Importantly: **voiced [b]** never observed in the "wrong" class

3-onset system: /pʰ/, /p/, /p~b/		
	high	low
*aspirated	pʰá	
*voiceless	pá	pʰà
*voiced		pà~bà

2-onset system: /pʰ/, /p/		
	high	low
*aspirated	pʰá	
*voiceless	pá	pà
*voiced		pà

Two onset consonant systems

Both tonal *and* non-tonal (merged) speakers found with each

- Importantly: **voiced [b]** never observed in the "wrong" class

3-onset system: /pʰ/, /p/, /p~b/		
	high	low
*aspirated	pʰá	
*voiceless	pá	pʰà
*voiced		pà~bà

2-onset system: /pʰ/, /p/		
	high	low
*aspirated	pʰá	
*voiceless	pá	pà
*voiced		pà

Check the corpus

The interview recordings

- Recordings of 19 Diaspora-raised speakers
  - Do they also vary in whether or not they use tone?
    - Yes: 11/19 have higher pitch in H than L words
  - Do they also show consistent prevoicing behavior?
    - Yes: optional voicing only on the same set
    - Fricatives as well as stops

Summary

- Lots of variation
  - presence/absence of tone
  - two consonant systems
- What remains stable?
  - C-V timing
  - Categories distinguished by tone

Roadmap

- Introduction
- Articulatory Coordination
- **Studies**
  - Tibetan tone
  - **English diphthongs**
- Conclusion

Measuring to modeling

(Haken et al. 1985, Saltzman & Munhall 1989, Nam & Saltzman 2003)

- Model kinematics as critically-damped mass-spring oscillator
- In-phase/anti-phase/etc. determine relative timing

$$ma + bv + k(x - C) = 0$$

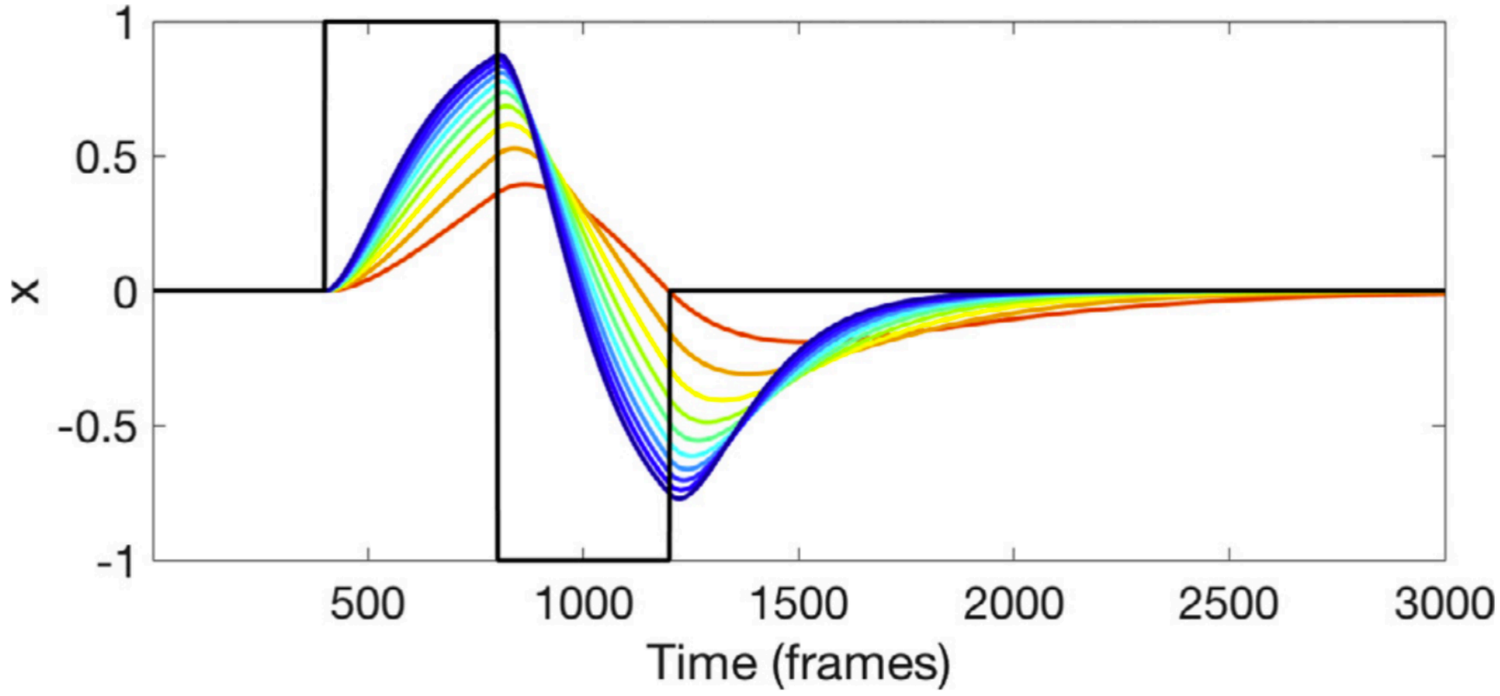
acceleration

velocity

position

stiffness

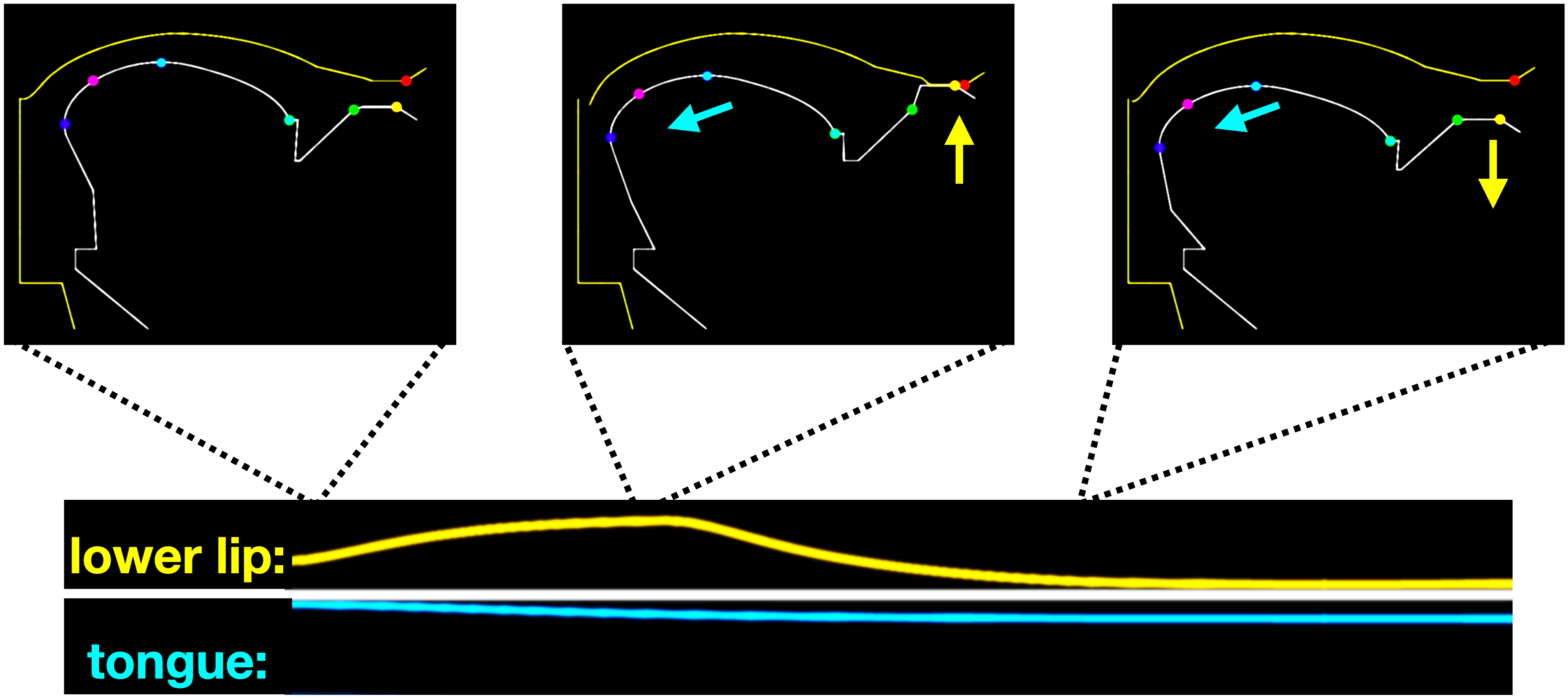
target



(Iskarous 2013)

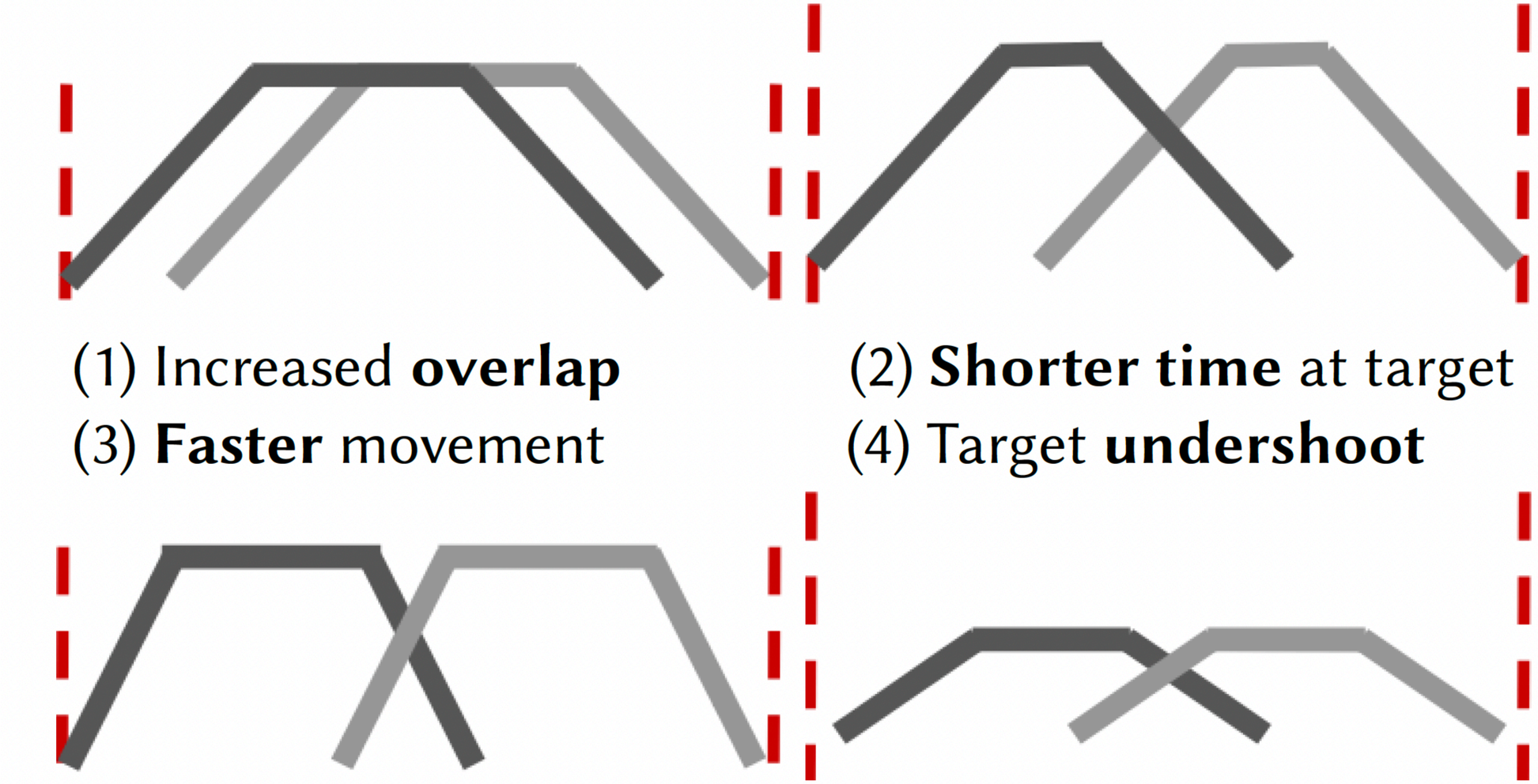
Articulatory simulation

TADA: Task Dynamics Application *(Nam et al. 2004)*



Simulating reduction

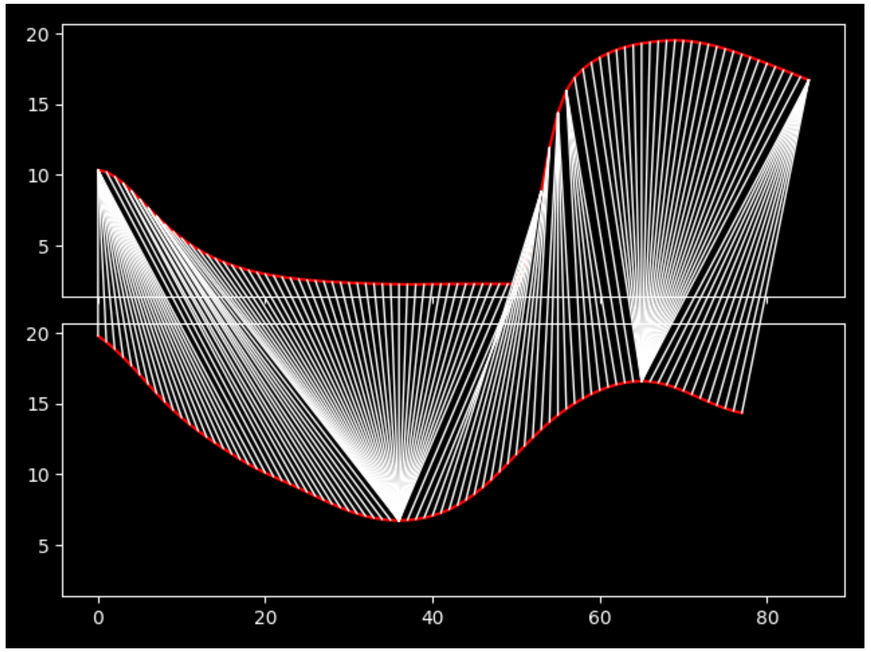
probabilistic or otherwise

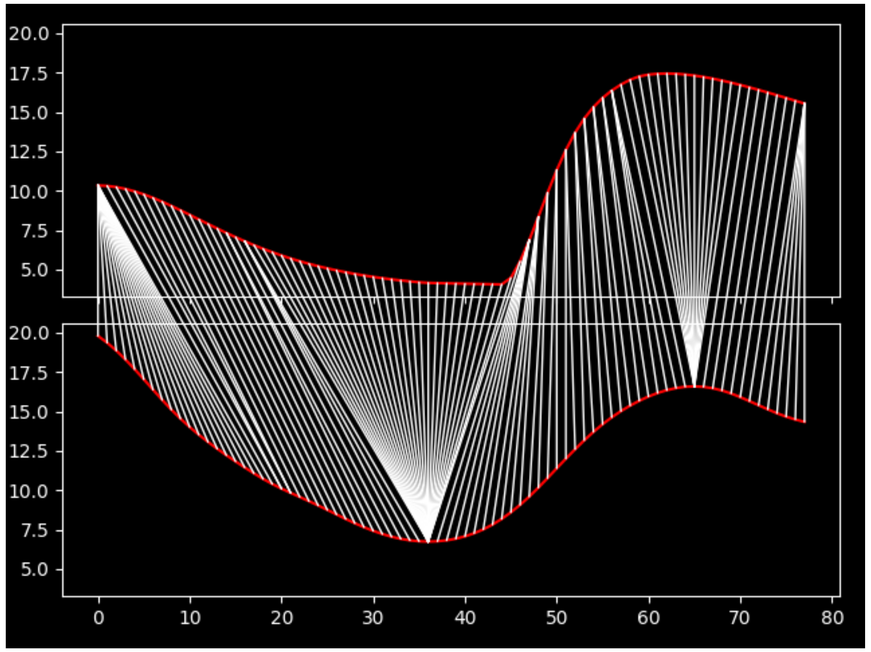


<five> study

- why? independent articulators (lips, tongue dorsum)
- simulate all permutations of four reduction modes, compare to 525 tokens from X-ray Microbeam Database
- Find best-fit curves with DTW; correlations with duration

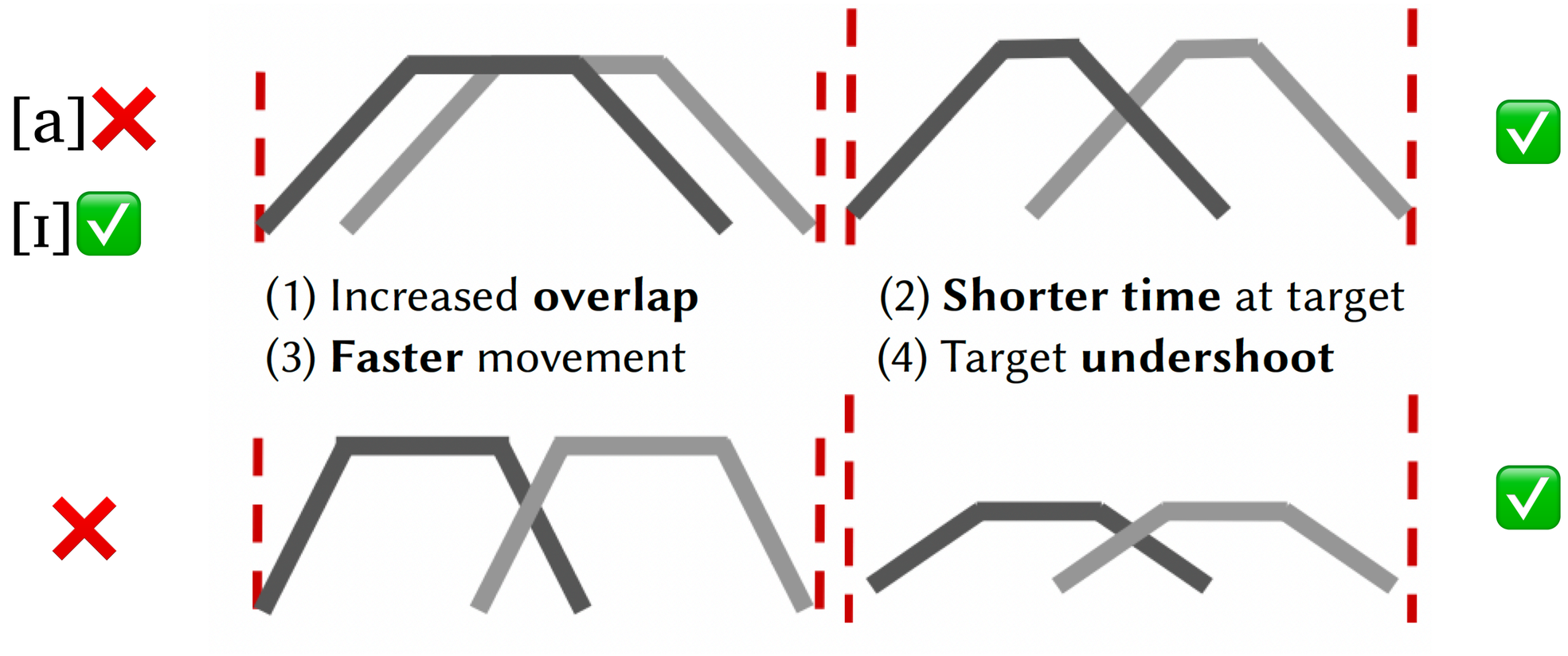
Bad fit





Good fit

Simulation results



- However, only (2) time at target was correlated with duration
- Correlations among (1)-(2)-(4) suggest role in defining shape

Interpretation

in <five>

- Overlap from [ɪ] but not from [a] suggests:
  - [a] portion of dipthong timed to rest of word
  - [ɪ] portion more free to vary across tokens
- Stiffness may be invariant
- Other factors can vary across tokens

Roadmap

- Introduction
- Articulatory Coordination
- Studies
  - Tibetan tone
  - English diphthongs
- **Conclusion**

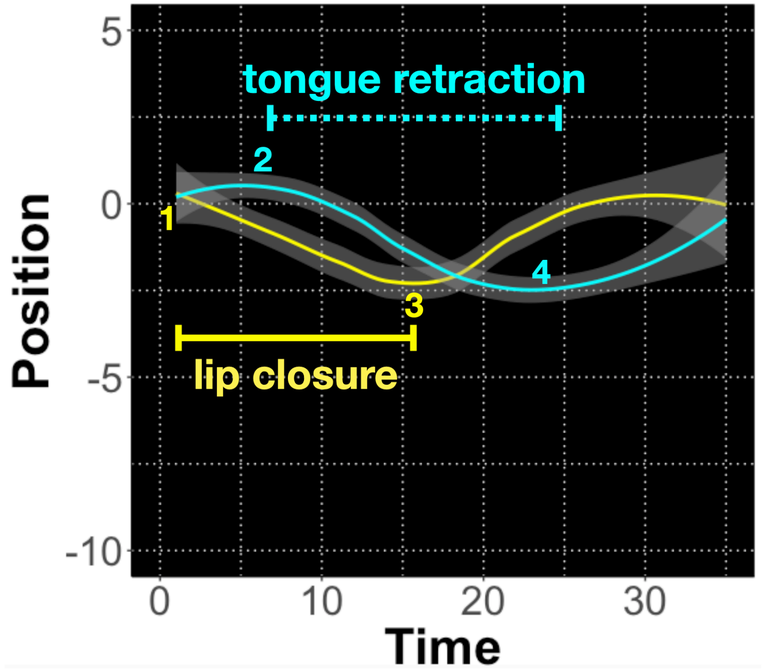
Summary

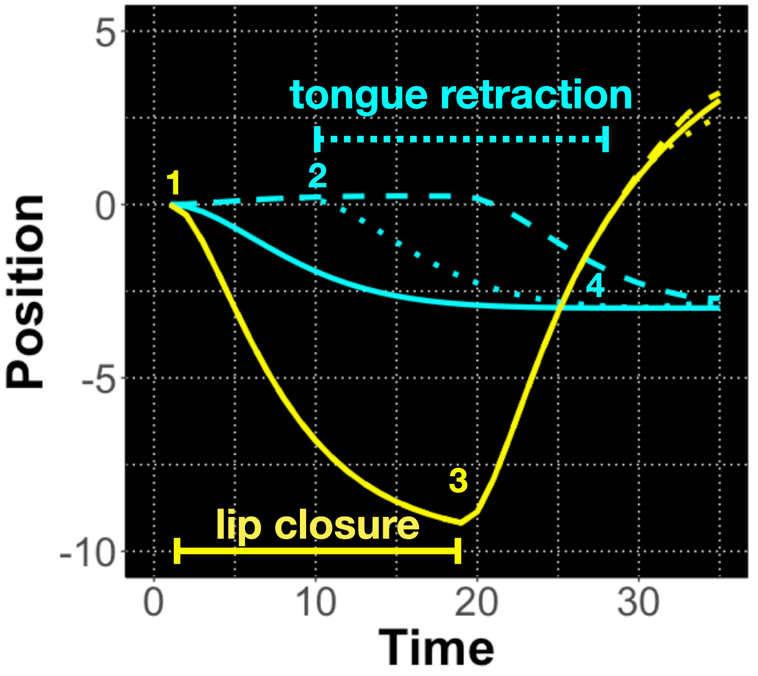
- Tibetan tone:
  - C-V timing suggestive of tone gesture… even when neutralized
  - Persistance of tone category in relationship with consonants
- English diphthongs:
  - stable timing [a], not [ɪ]—perhaps phonological
  - intergestural coordination more stable than gesture duration or spatial target

Next steps

beyond <five>

- Simulating Tibetan tone
- Level vs. sloped tone targets?
- Simulation & comparison of more data
  - Articulatory annotation?
- Deeper into Tibetan acoustics





Conclusions

- Tone relates to timing… but it's complicated
- Diphthongs appear onset-coordinted
- Relative timing looks more like phonology than other articulatory parameters do
- Methods:
  - Variation → look for stability and correlations
  - Modeling to disentangle articulation-acoustics mapping

References

- Browman, Catherine P. & Louis Goldstein. 1988. Some Notes on Syllable Structure in Articulatory Phonology. *Phonetica* 45(2–4). 140–155. https://doi.org/10.1159/000261823.
- Gao, Man. 2008. Tonal alignment in Mandarin Chinese: An articulatory phonology account. *Unpublished Doctoral Dissertation (Linguistics), Yale University, CT*.
- Geissler, Christopher. 2021. *Temporal articulatory stability, phonological variation, and lexical contrast preservation in diaspora Tibetan*. Yale University. https://elischolar.library.yale.edu/gsas_dissertations/52.
- Geissler, Christopher, Jason Shaw, Fang Hu & Mark Tiede. 2021. Eccentric CV timing across speakers of diaspora Tibetan with and without lexical tone contrasts. In *Proceedings of the 12th International Seminar on Speech Production*.
- Goldstein, Louis, Dani Byrd & Elliot Saltzman. 2006. The role of vocal tract gestural action units in understanding the evolution of phonology. In Michael A. Arbib (ed.), *Action to Language via the Mirror Neuron System*, 215–249. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511541599.008.
- Goldstein, Louis, Hosung Nam, Elliot Saltzman & Ioana Chitoran. 2009. Coupled oscillator planning model of speech timing and syllable structure. *Frontiers in phonetics and speech science* 239–250.
- Haken, Hermann, JA Scott Kelso & Heinz Bunz. 1985. A theoretical model of phase transitions in human hand movements. *Biological cybernetics* 51(5). 347–356.
- Iskarous, Khalil, Christine Mooshammer, Phil Hoole, Daniel Recasens, Christine H. Shadle, Elliot Saltzman & D. H. Whalen. 2013. The coarticulation/invariance scale: Mutual information as a measure of coarticulation resistance, motor synergy, and articulatory invariance. *The Journal of the Acoustical Society of America* 134(2). 1271–1282. https://doi.org/10.1121/1.4812855.
- Karlin, Robin. 2014. The articulatory TBU: gestural coordination of tone in Thai. In *Cornell Working Papers in Linguistics*.
- Nam, Hosung & Elliot Saltzman. 2003. A competitive, coupled oscillator model of syllable structure. In *Proceedings ofthe 15th International Congress of the Phonetic Sciences*.
- Niemann, Henrik, Doris Mücke, Hosung Nam, Louis Goldstein & Martine Grice. 2011. Tones as Gestures: The Case of Italian and German. In *Proceedings of the 17th International Congress of the Phonetic Sciences*, 1486–1489.
- Saltzman, Elliot & J. A. Kelso. 1987. Skilled actions: A task-dynamic approach. *Psychological Review* 94(1). 84–106. https://doi.org/10.1037/0033-295X.94.1.84.
- Tiede, Mark. 2005. Mview: software for visualization and analysis of concurrently recorded movement data.
- Zhang, Muye, Christopher Geissler & Jason Shaw. 2019. Gestural representations of tone in Mandarin: Evidence from timing alternations. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Melbourne, Australia.

ཐུགས་རྗེ་ཆེ།

Thank you!

Vielen Dank!

Articulatory study

Geissler et al. (2021), Geissler (2021ch4)

- H1: variation in timing conditioned by presence/absence of lexical tone
  - speakers with tone contrast will have competitive coupling (pos. C-V lag)
  - speakers without tone contrast will have in-phase C-V timing (no C-V lag)
- H2: timing convergence:
  - all speakers will have similar coordination patterns despite interspeaker variation in presence/absence of tone
- What kind of tone contrast is there?
  - If H-∅, then difference will be visible in high vs. low tone words
  - If H-L, then no difference in timing by tone.

EMA Study conclusions

- H1: variation in timing conditioned by presence/absence of lexical tone
  - speakers with tone contrast will have competitive coupling (pos. C-V lag)
  - speakers without tone contrast will have in-phase C-V timing (no C-V lag)
- ✔ **H2: timing convergence:**
  - all speakers have similar coordination patterns despite interspeaker variation in presence/absence of tone
- What kind of tone contrast is there?
  - If H-∅, then difference will be visible in high vs. low tone words
  - ✔ If **H-L**, then no difference in timing by tone.

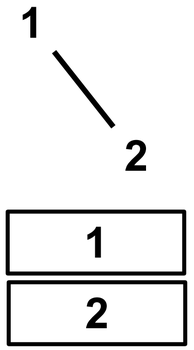
The temporal basis of complex segments

Shaw et al. 2019

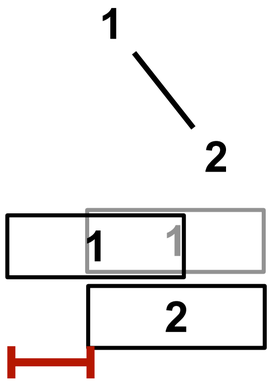
The temporal basis of complex segments

Shaw (2019): predictions

In-phase



In-phase + lag (offset)



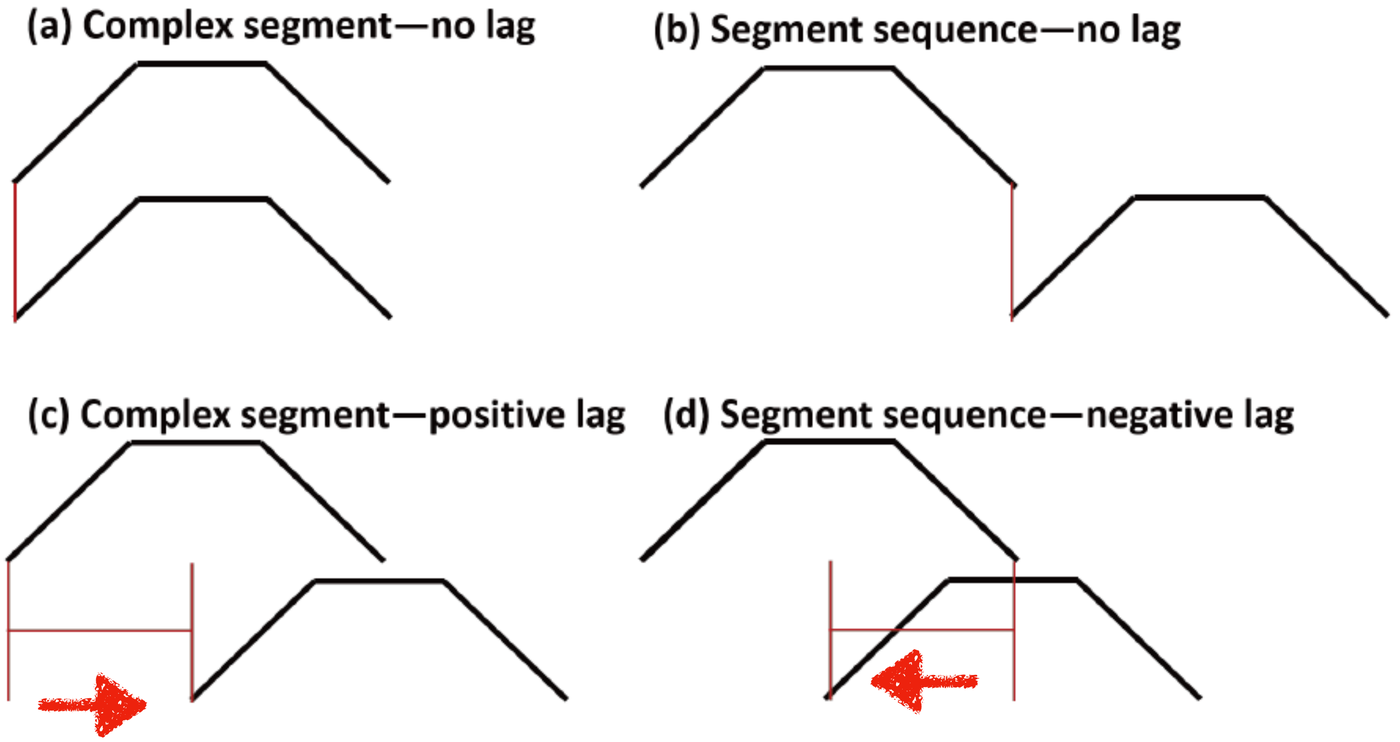
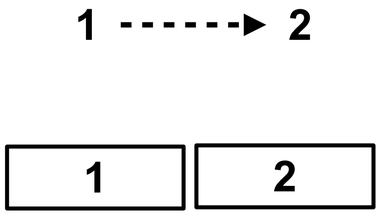
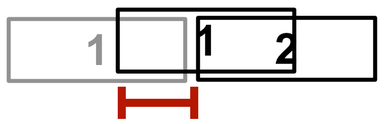


Figure 1: Hypothesized gestural coordination patterns for complex segments (a), (c) and segment sequences (b), (d)

Anti-Phase



Anti-Phase - lag (offset)



The temporal basis of complex segments

Shaw (2019): results

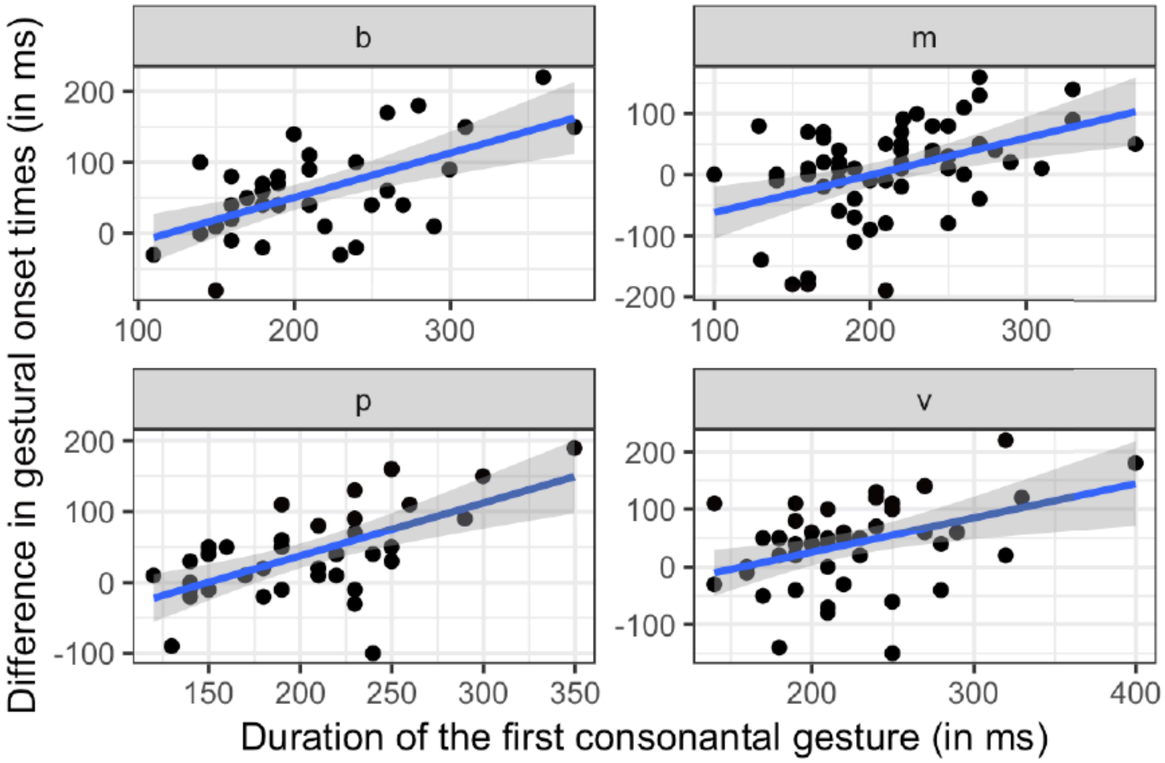


Figure 4: Correlations for the data from the English experiment

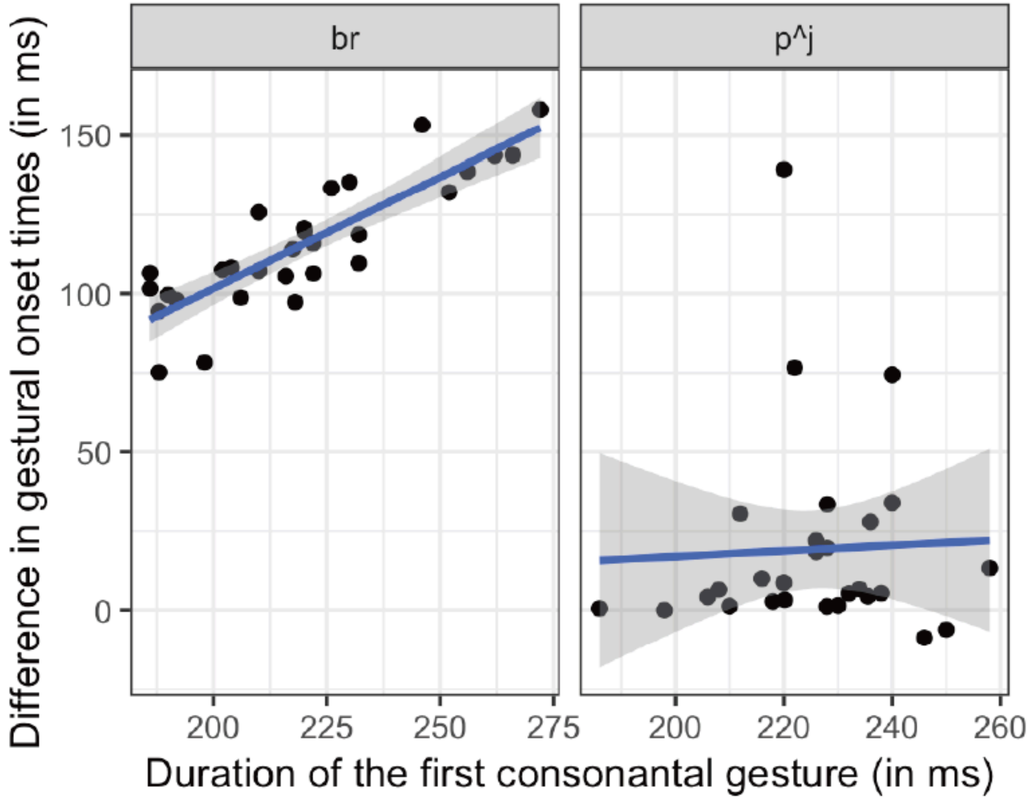


Figure 2: Correlations for the Russian data

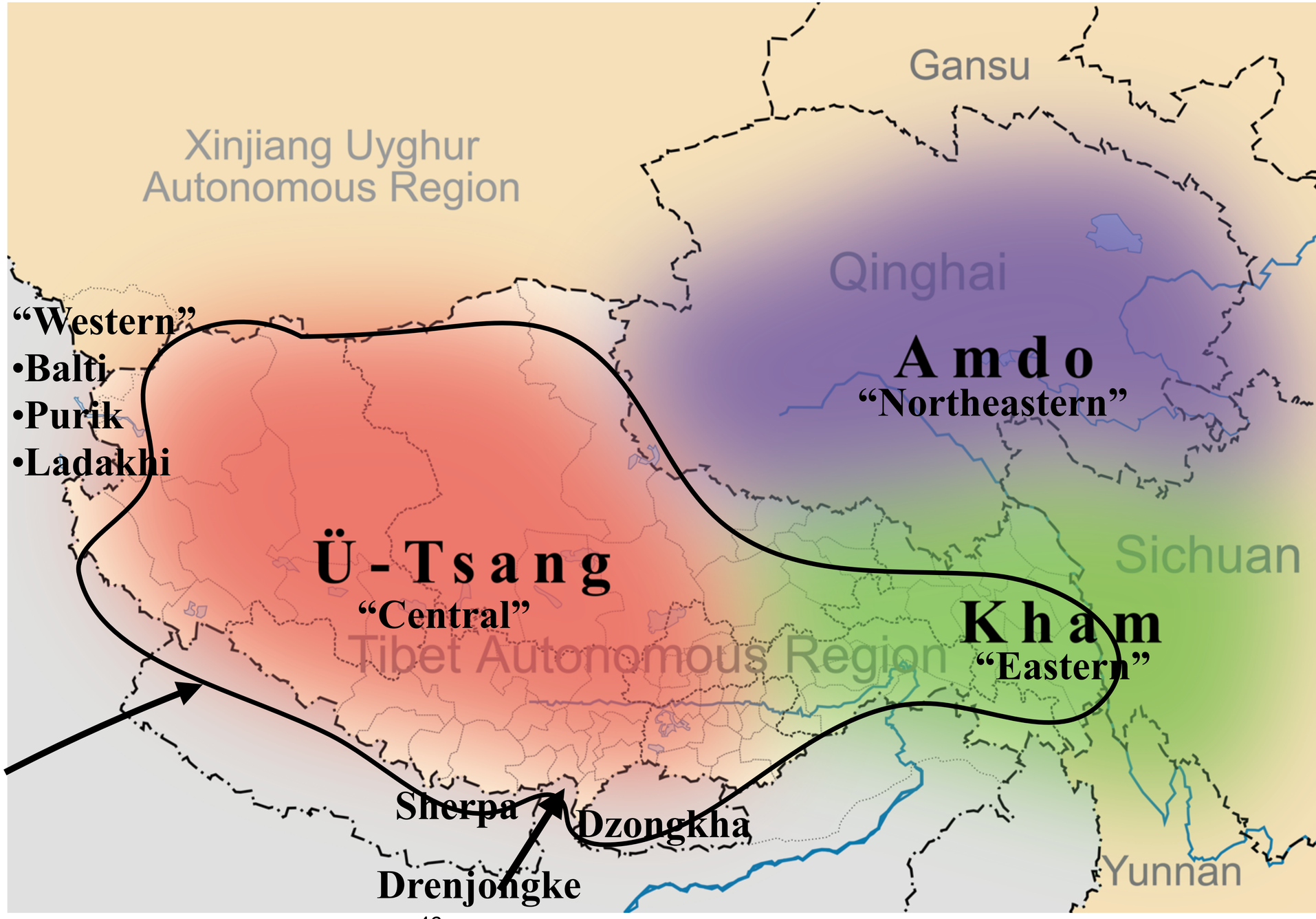
Tibetan dialects

Tibetan

བོད་སྐད་

- “archaic”/ “cluster”
- “innovative”/ “non-cluster”
- dialect continuum
- post-1959 diaspora

Approx. extent of tone



Dialects: Natural laboratory

- tonogenesis
- laryngeal variation
- cluster simplification
- vowel shifts, spirantization, retroflexion, palatalization
- evidential, honorifics, modality, etc.

Written (Classical) Tibetan	Balti (Western)	Rebkong (Northeastern)	Tokpe Gola (Central)	Gloss
khrag	[kʂʌk]	[t͡ɕɤɣ]	[t^{h}ʌ́k] ([t^{h}ák])	‘blood’
rtswa	[xst͡soa]	[xt͡sa]	[t͡sá]	‘grass’
spyang ki	[spjaŋ.ˈku]	[xt͡ɕaŋ.ˈk^{h}ɤ]	[t͡ʃáŋ.gú]	‘wolf’
bcu bdun	[t͡ɕub.ˈdʊn]	[t͡ɕɤb.ˈdɤn]	[t͡ʃúp.tő] ([t͡ʃúp.tỹ])	‘seventeen’

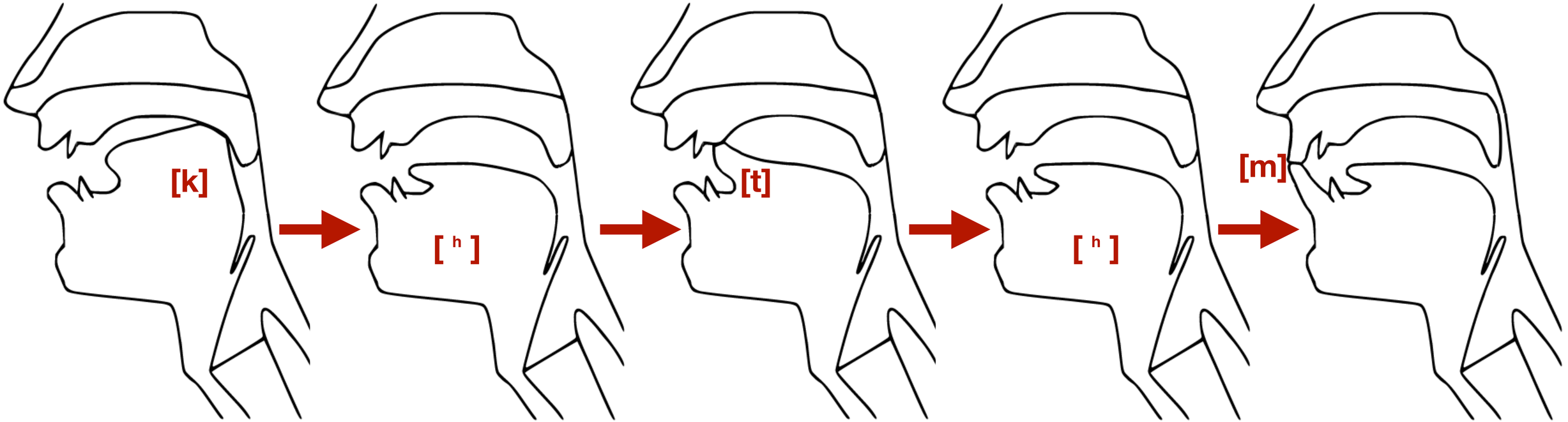
(Adapted from Caplow 2013)

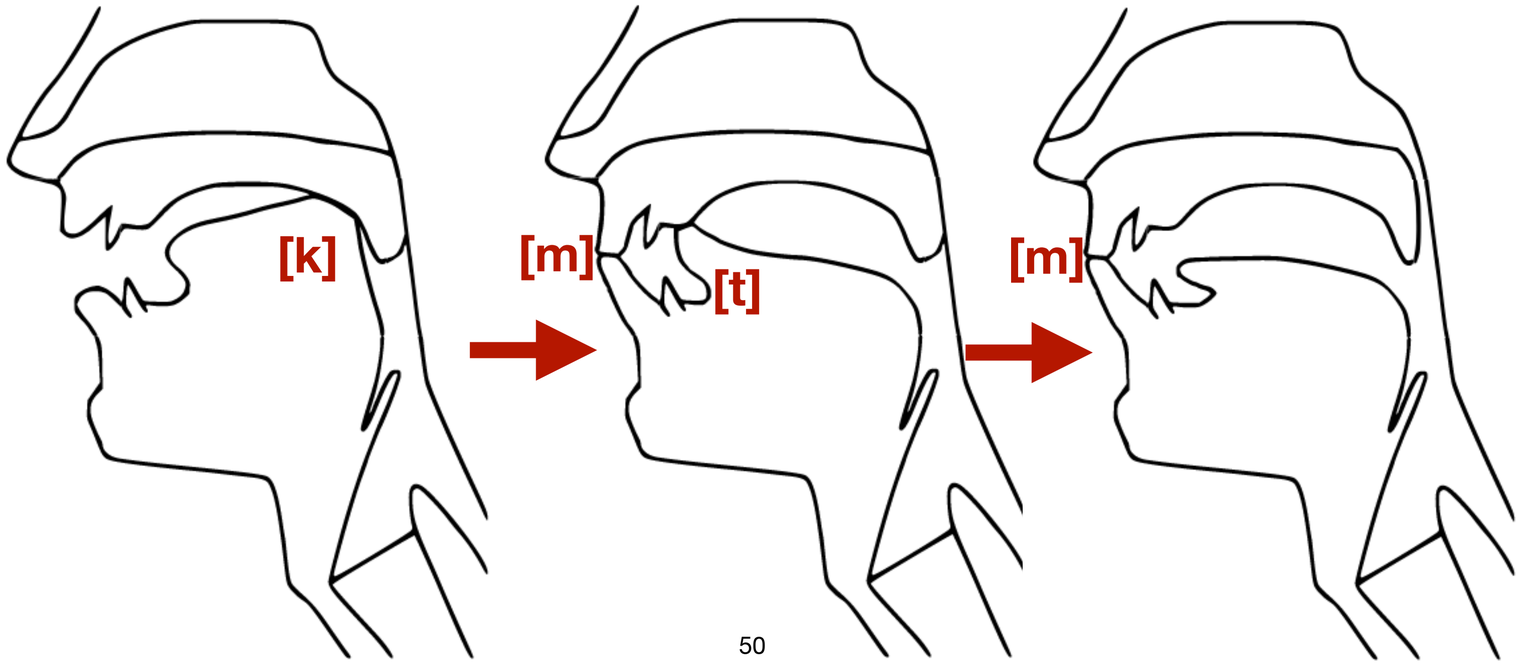
Tonogenesis
(tonal dialects only)

- Voiceless onsets > high tone
  - *p^{h}ar ‘over there’ > H
    *sa ‘earth’ > H
- Voiced onsets > low tone
  - *bar ‘between’ > L
    *za ‘eat’ > L
    *mar ‘butter’ > L
- Sonorants with pre-initial > high tone
  - *sman ‘medicine’ > H

Laryngeal contrasts

	Etymological onsets							Innovative features
Orthography	སྤ་	ཕ་	བ་	སྦ་	ས་	ཟ་	བཟ་	
Old Tibetan	sᵊpa	pʰa	ba	sᵊba	sa	za	bᵊza	aspiration allphonic
Northeastern and Western dialects	spa	pʰa	ba ~ wa	ʁba	sa	za	za	cluster simplification aspirated/unaspirated contrast
Eastern dialects	pá	pʰá	pà	bà	sá	zà	zà	tonogenesis cluster simplification
Central dialects (Lhasa)	pá	pʰá	pʰà	pà	sá	sà	sà	voiced clusters > voiceless voiced simplex > aspirated



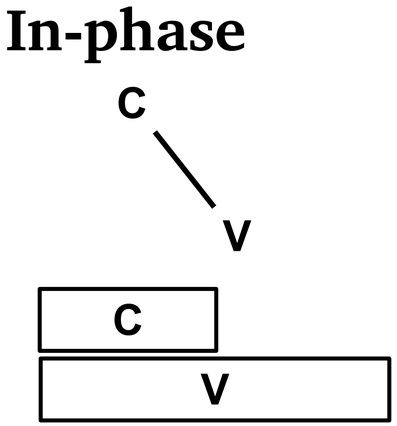


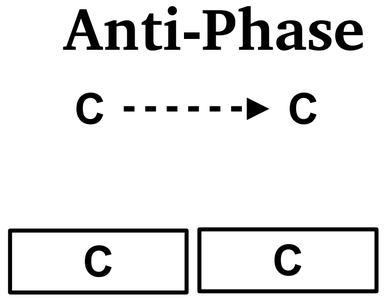
back to slide 7

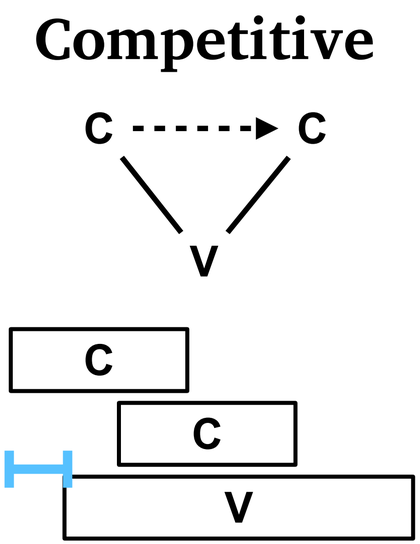
Interactive Sagittal Section

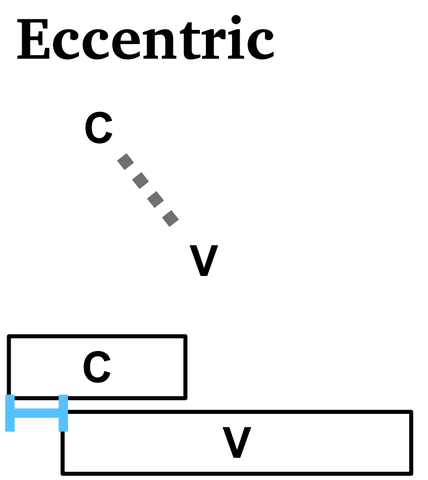
Coordinating gestures in time

- Gestural coupling modes:
  - *In-phase coupling:* (synchronous) and *Anti-phase coupling* (sequential) are most stable
  - *Competitive coupling:* combination of in-phase and anti-phase coupling relations
  - *Eccentric coupling:* one coupling relation, just not intrinsically stable







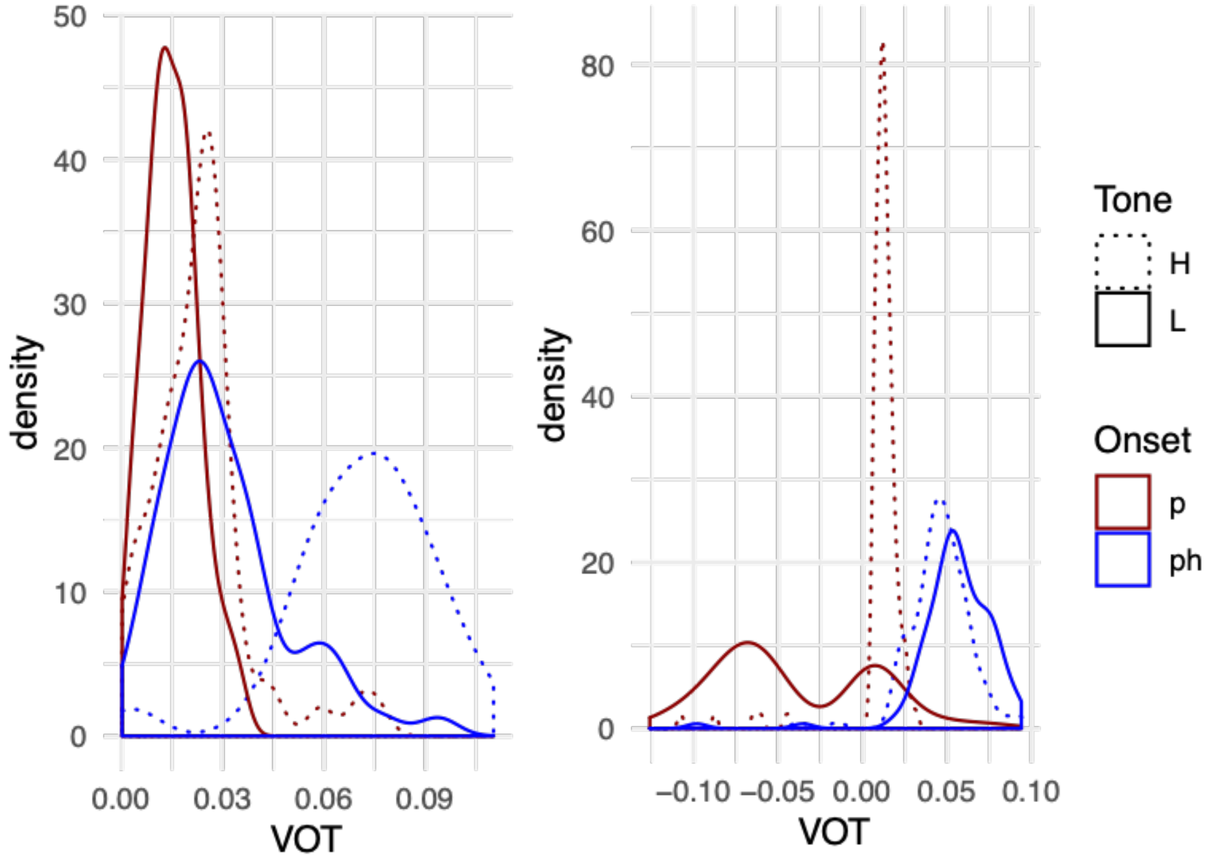


(Nam & Saltzman 2003, Nam et al. 2009, Goldstein 2011)

Two systems of laryngeal contrasts

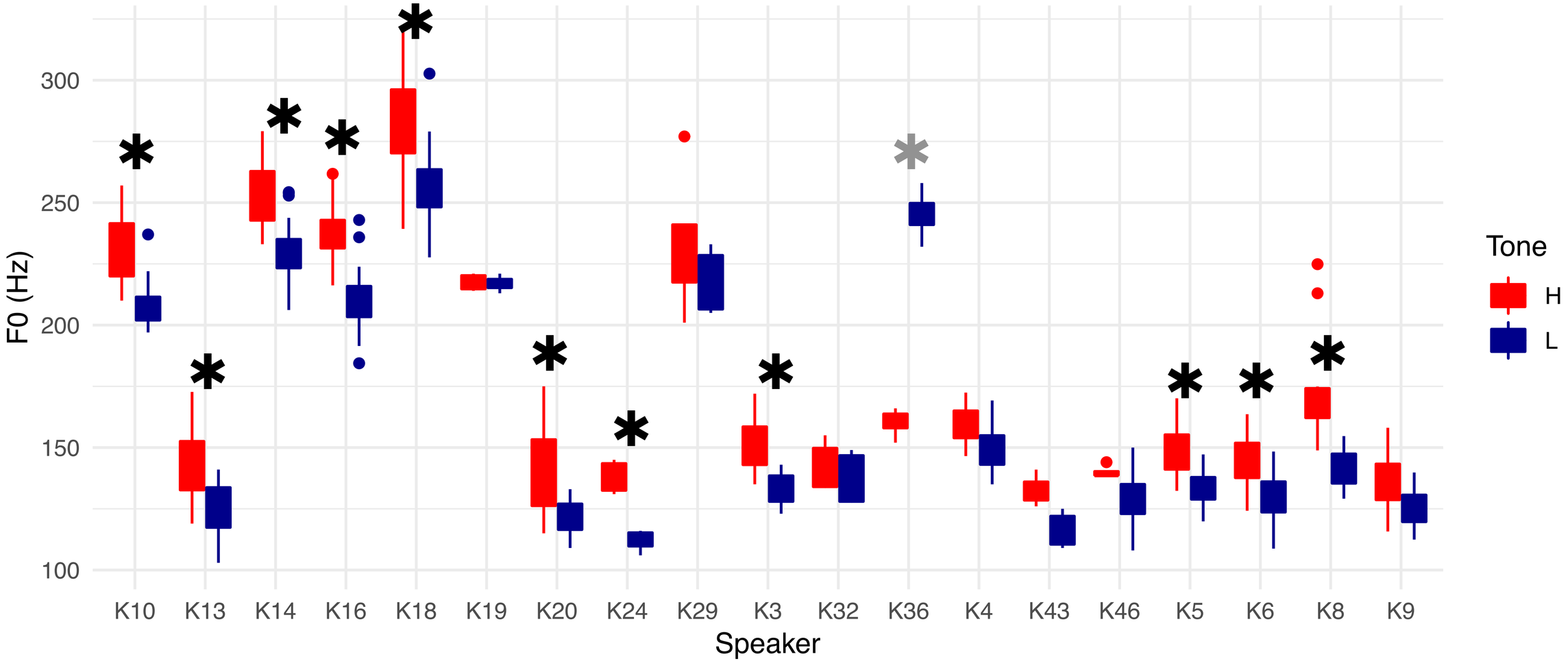
even in speakers with no F0 contrast (!!!)

- Both conditioned by etymological tone category:
- Left speaker
  - no prevoicing
  - long VOT only with H tone
- Right speaker:
  - prevoicing with L tone
  - long VOT with both tones

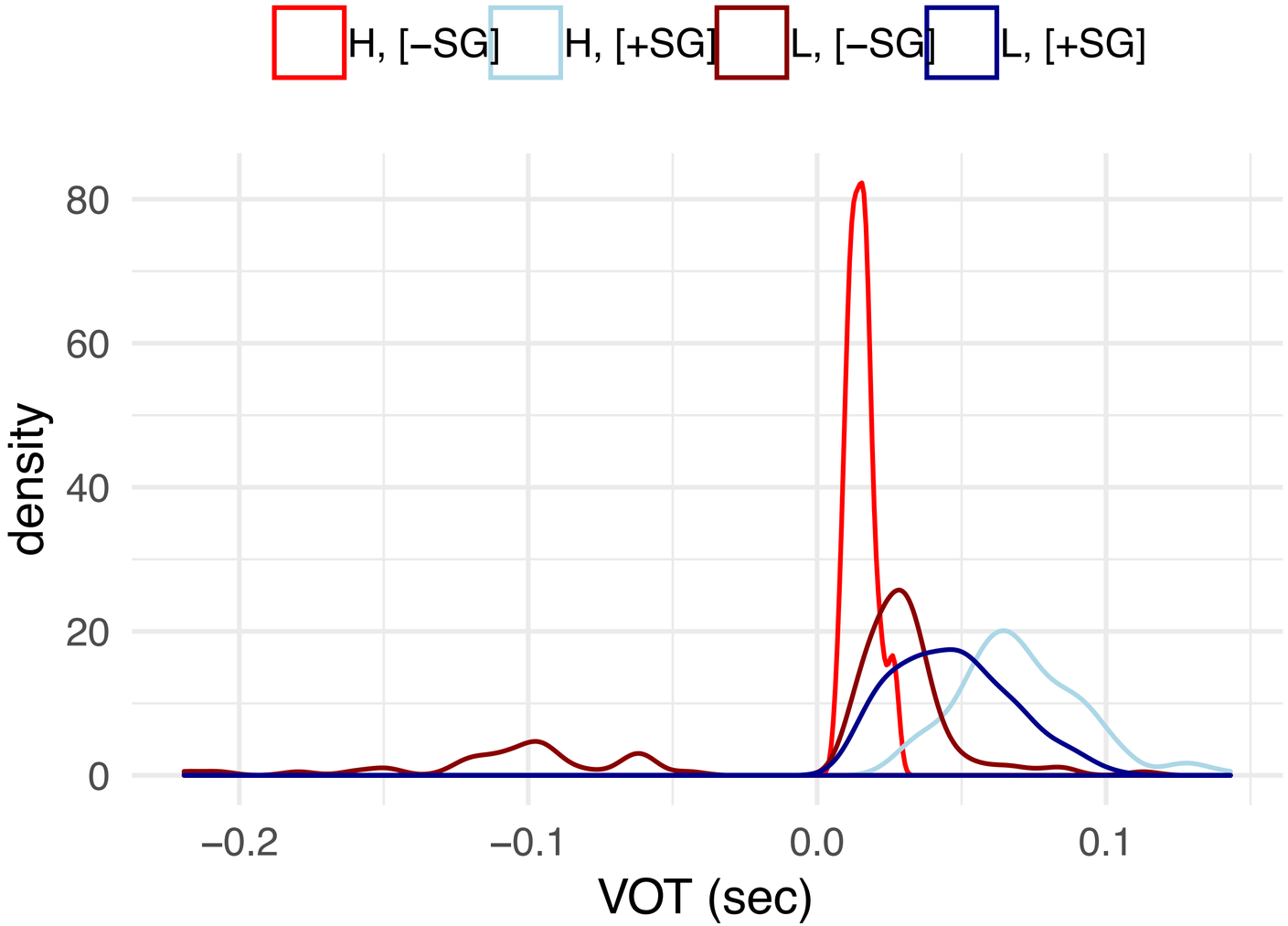


Does H have higher pitch than L?

Yes for 11/19, no for 7/19



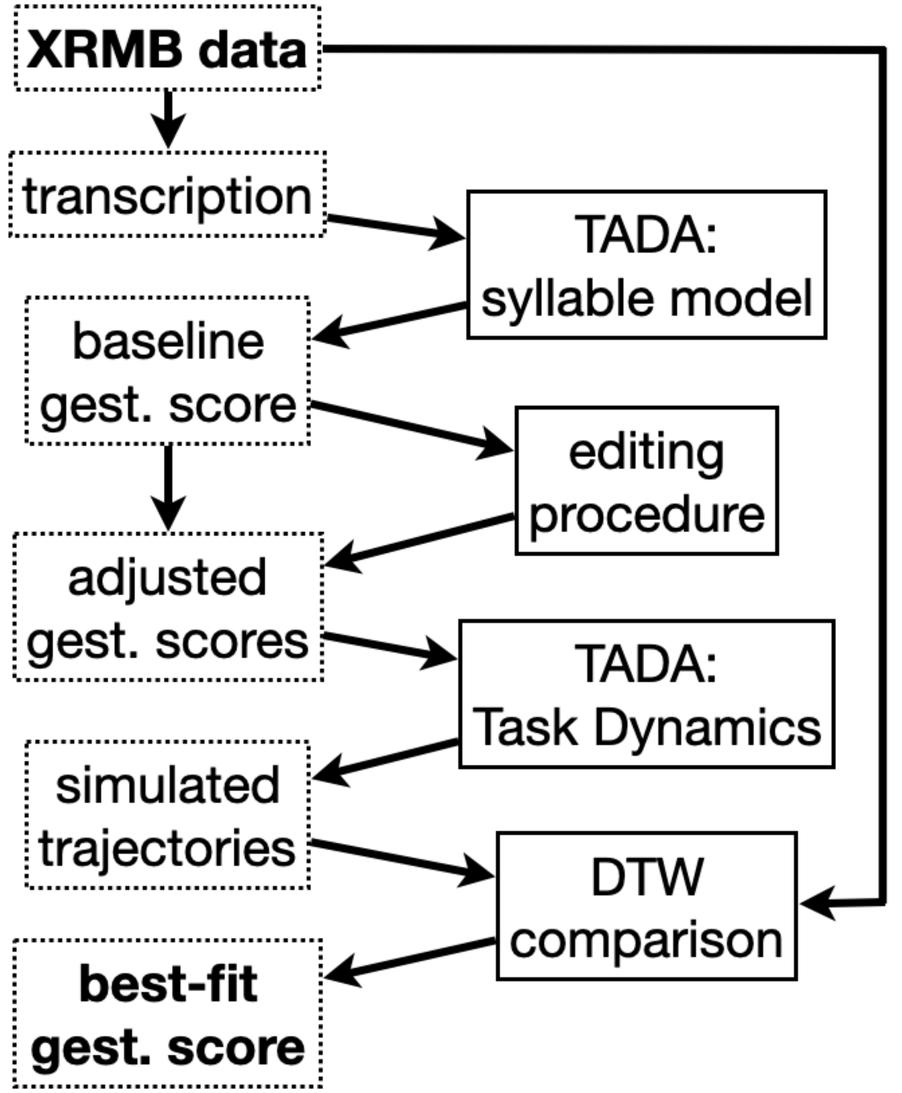
Consonant and tone categories



<five> study: methods

O'Reilly, Geissler, & Tang (2023)

- Ideal test case?
  - diphthongs: all four modes
  - C's with lips, V's with tongue
  - available data



Timing in phonology and/or phonetics?

- “Discrete Phonology” vs. “Gradient Phonetics”
- Speech timing as phonology
  - Is timing *intrinsic* or *extrinsic* to phonology?
  - Are gestures coordinated at *beginning* or *end*?
  - *Symbolic* vs. *phonetically-enriched* representations?

(Turk & Shattuck-Hufnagel 2021)