

From phonetic to phonological time

Christopher Geissler

Department of Linguistics

Carleton College

April 15, 2024

Slides available on cageissler.github.io/resources

Roadmap

- Biography
- Phonology, phonetics, and time
- Types of evidence
 - Intergestural timing—Tibetan tonality
 - Duration tradeoffs—Northern Saami
 - Simulated trajectories
- Conclusion

The Swarthmore years...

- Swarthmore '13: Linguistics, Religion
- Study abroad: Tibetan Studies Semester (Dharamsala, India)
 - Research project: dialect contact, or religious ethnography?
- Summer with Living Tongues dictionary of Koro
- Thesis: “Towards a phonetic description of Koro”

... a bizarrely linear doctoral program...

- LSA Institute @ UMich: Khalil Iskarous' Articulatory Phonology
- NYC's Endangered Language Alliance for 1 year
- Yale Linguistics!
 - incoming interests: phonetics, phonology, historical ling
 - fieldwork in Nepal—dialect contact, based on my paper abroad
 - Dissertation combined fieldwork with laboratory experiments

... still employed, so far...

- 2021-2023 @ Heinrich Heine U Düsseldorf
 - teaching, setting up new lab, not getting grants
- 2023-2024 @ Carleton
 - teaching, collaborating with students
- ... TBA?

...unsolicited advice.

- ... Ph.D.?
 - IFF you want to *be in grad school*
- ... academic career?
 - IFF able to move, internationally, every year, for years
- ... regrets @ Swat?
 - dropping Chinese; not taking stats; thesis topic
 - honestly wouldn't do differently. Except the thesis.

Roadmap

- Biography
- Phonology, phonetics, and time
- Types of evidence
 - Intergestural timing—Tibetan tonality
 - Duration tradeoffs—Northern Saami
 - Simulated trajectories
- Conclusion

Phonology: basic Categorical behavior

- In German, voiced consonants are voiceless when they occur at the end of words (but not elsewhere):
 - *Maus* ‘mouse’ [maʊs], but plural *Mäuse* [mɔʏzə]
 - *Rad* ‘wheel’ [rat], but plural *Räder* [rɛdə]
 - compare:
Rat ‘council’ [rat], but plural *Räte* [rɛtə]

Phonology: basic Categorical behavior

- In German, voiced consonants are voiceless when they occur at the end of words (but not elsewhere):
 - *Maus* ‘mouse’ [maʊs], but plural *Mäuse* [mɔʏzə]
 - *Rad* ‘wheel’ [rat], but plural *Räder* [rɛdə]
 - compare:
Rat ‘council’ [rat], but plural *Räte* [rɛtə]

Linguists are really good at this

Phonology: advanced

Probabilistic behavior

- In English, t/d at the end of a word sometimes isn't there
 - *rift* = [ɹɪft] or [ɹɪf_]; *build* = [bɪld] or [bɪł]
 - More likely among some groups
 - More likely in some social contexts
 - More likely around some sounds
 - More likely in *mist* than in *missed*

Phonology: advanced Probabilistic behavior

- In English, t/d at the end of a word sometimes isn't there
 - *rift* = [ɹɪft] or [ɹɪf_]; *build* = [bɪld] or [bɪł]
 - More likely among some groups
 - More likely in some social contexts
 - More likely around some sounds
 - More likely in *mist* than in *missed*

Linguists get excited about this

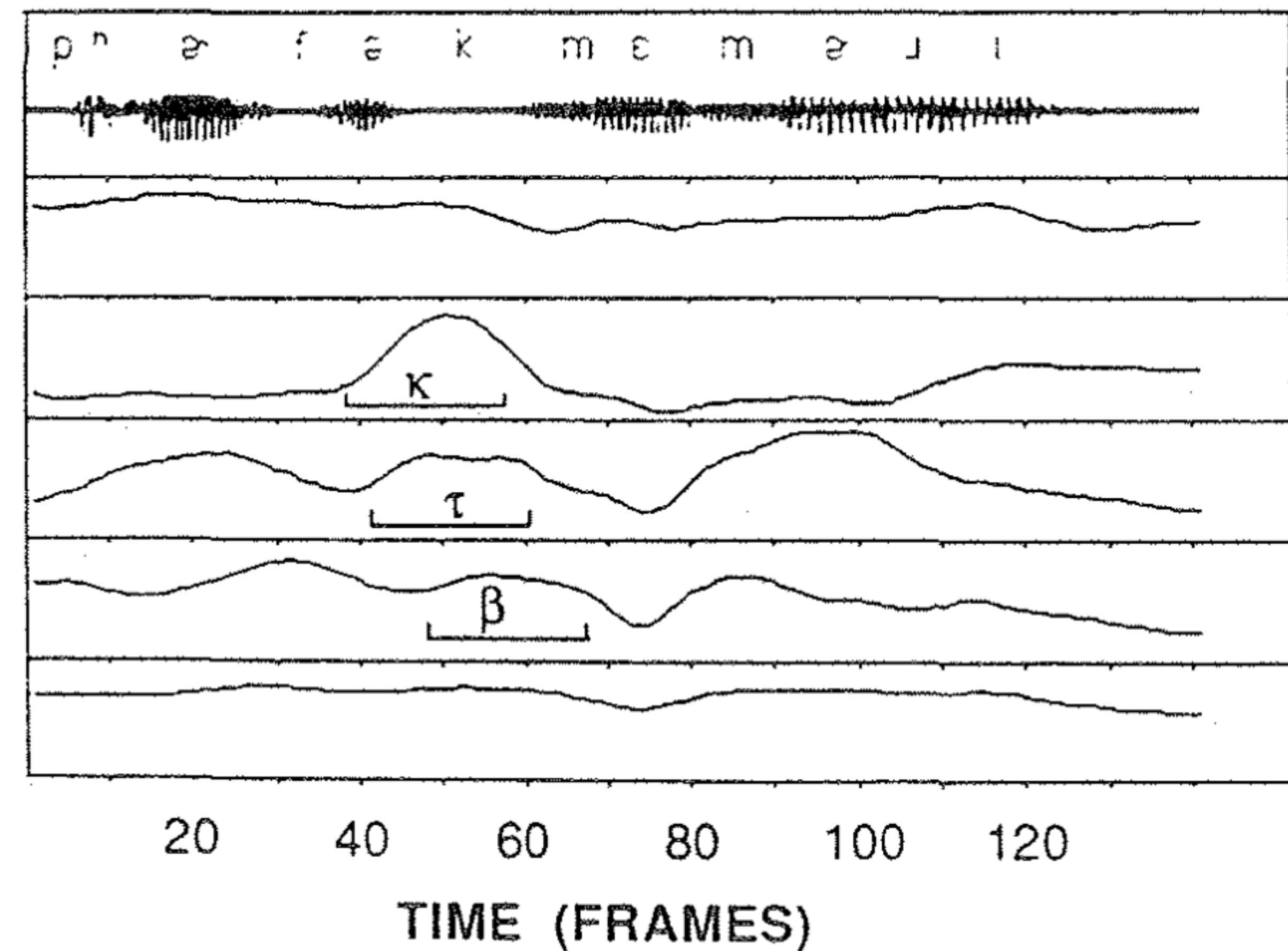
...uh-oh

- *Perfect memory*
- At least some “deleted” t’s/d’s are visible in articulation, but not in acoustics
- (Actually it’s most)

Midsagittal sections

(Browman & Goldstein 1988, Purse 2019)

AUDIO
WAVEFORM
VELUM
TONGUE
REAR
TONGUE
BLADE
LOWER
LIP
JAW



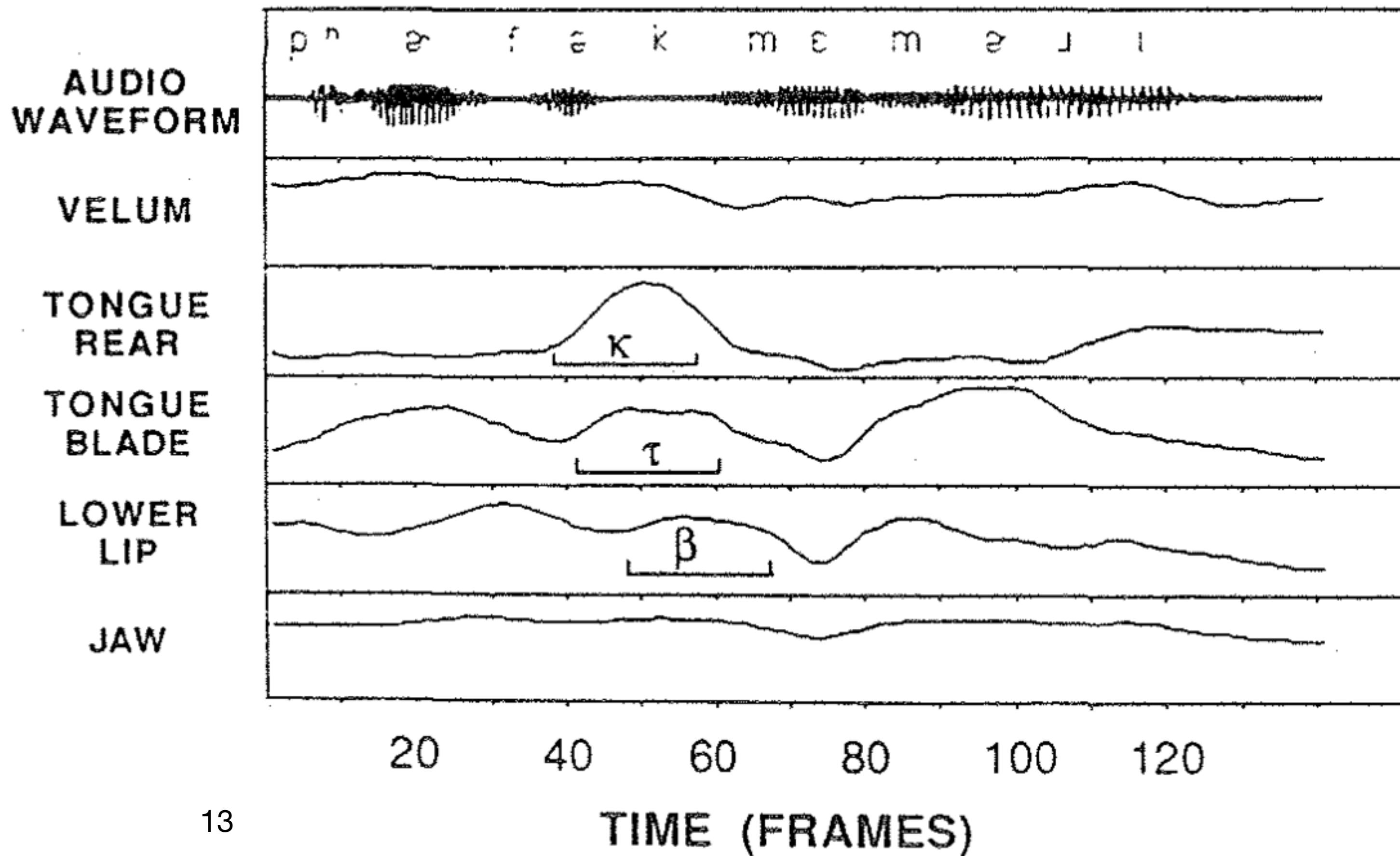
...uh-oh

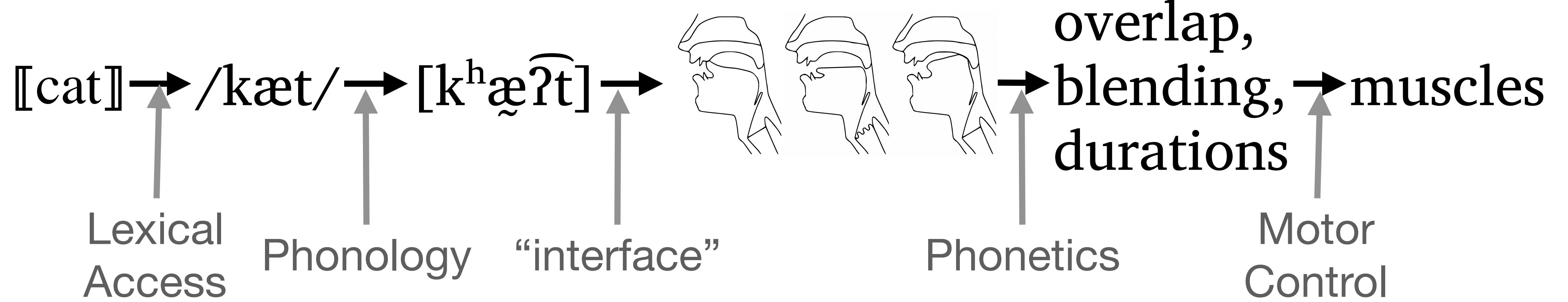
- *Perfect memory*
- At least some “deleted” t’s/d’s are visible in articulation, but not in acoustics
- (Actually it’s most)

Midsagittal sections

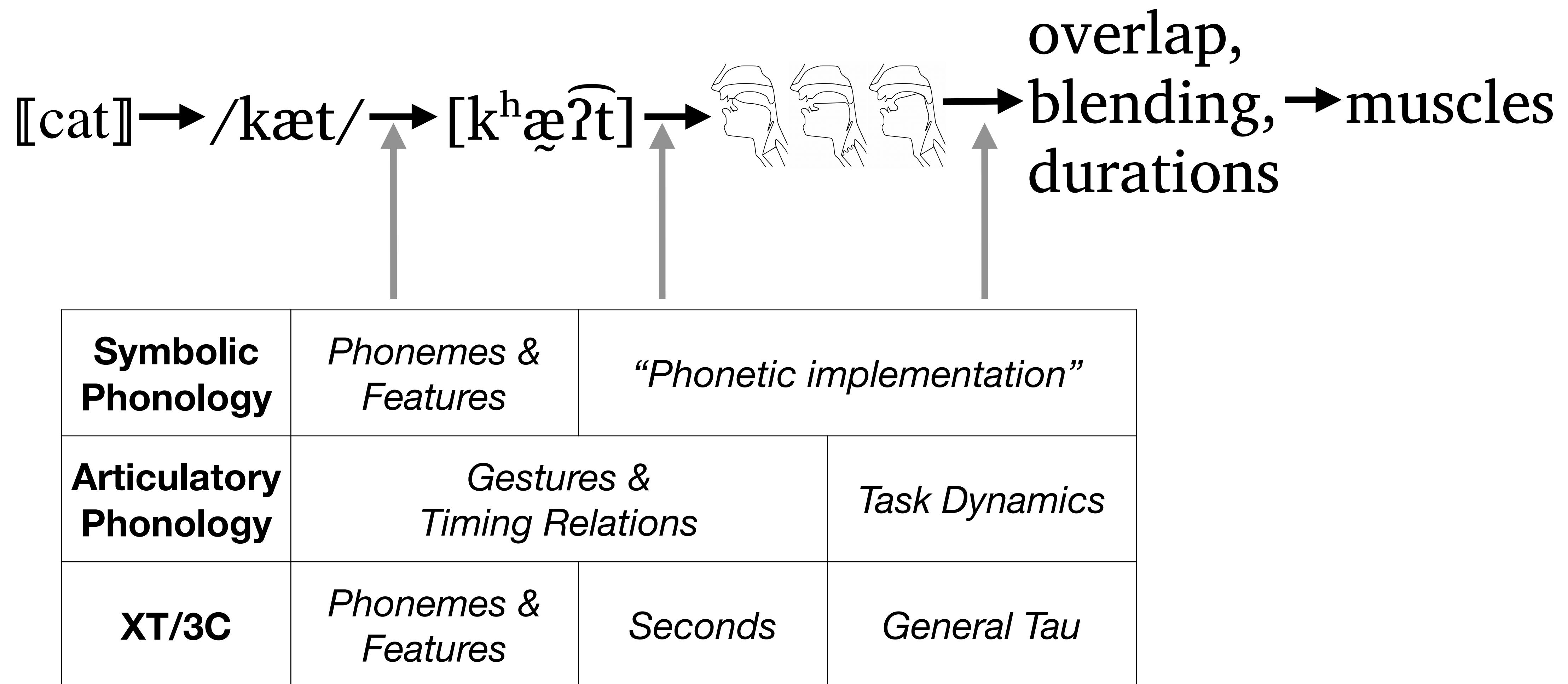
(Browman & Goldstein 1988, Purse 2019)

Gestures!
... but how are they coordinated?





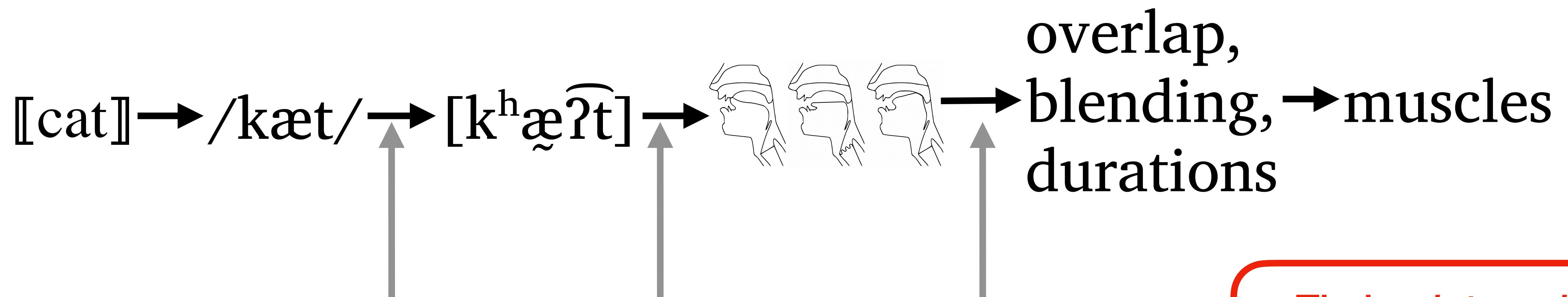
Representational units



AP: Browman & Goldstein (1986) et seq.; TD: Saltzman & Munhall (1989)

XT/3C: Turk & Shattuck-Hufnagel (2020); Tau: Lee (1998)

Representational units



Symbolic Phonology	<i>Phonemes & Features</i>	<i>“Phonetic implementation”</i>
Articulatory Phonology	<i>Gestures & Timing Relations</i>	<i>Task Dynamics</i>
XT/3C	<i>Phonemes & Features</i>	<i>Seconds</i>

Timing *internal* or
external to phonology?

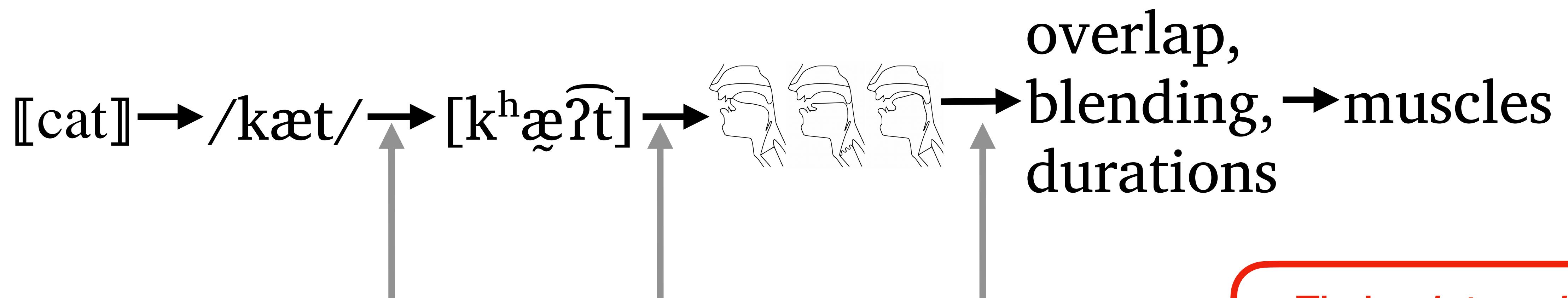
Both categorical and
continuous timing?

Which better fits
articulatory data?

Roadmap

- Biography
- Phonology, phonetics, and time
- Types of evidence
 - **Intergestural timing—Tibetan tonality**
 - Duration tradeoffs—Northern Saami
 - Simulated trajectories
- Conclusion

Representational units



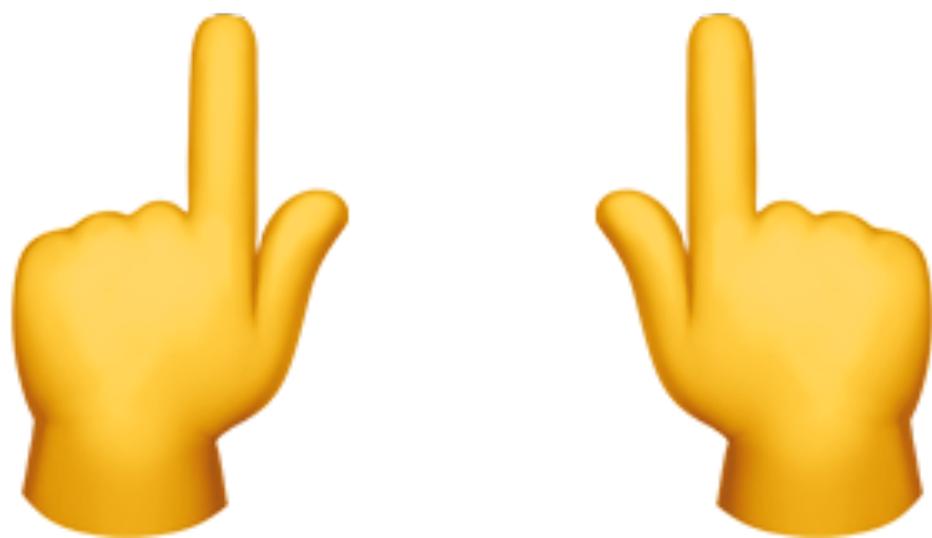
Symbolic Phonology	<i>Phonemes & Features</i>	<i>“Phonetic implementation”</i>
Articulatory Phonology	<i>Gestures & Timing Relations</i>	<i>Task Dynamics</i>
XT/3C	<i>Phonemes & Features</i>	<i>Seconds</i>

Timing *internal* or
external to phonology?

Both categorical and
continuous timing?

Which better fits
articulatory data?

*****Bimanual tapping interlude*****



Oscillators

- Synchronization in non-speech and speech movements:
 - “pa... pa... pa... pa.pa[...]pa.pa.pa.pa”
 - “ap... ap... ap... ap.ap.[...]pa.pa.pa.pa”
- Tapping: “in-phase” more stable than “anti-phase”
(both more stable than any other phasing)
... in speech too?

CV vs. VC syllables

in-phase

[pa]	
LIPS	Labial closure
TONGUE TIP	
TONGUE BODY	pharyngeal wide

anti-phase

[ap]	
LIPS	labial closure
TONGUE TIP	
TONGUE BODY	pharyngeal wide

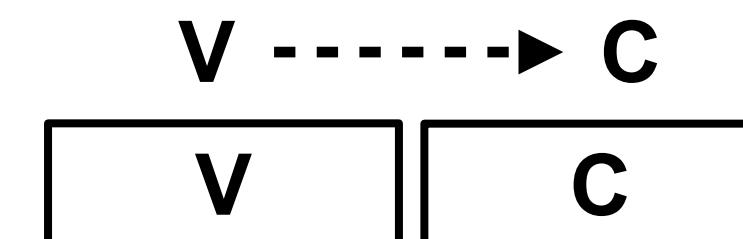
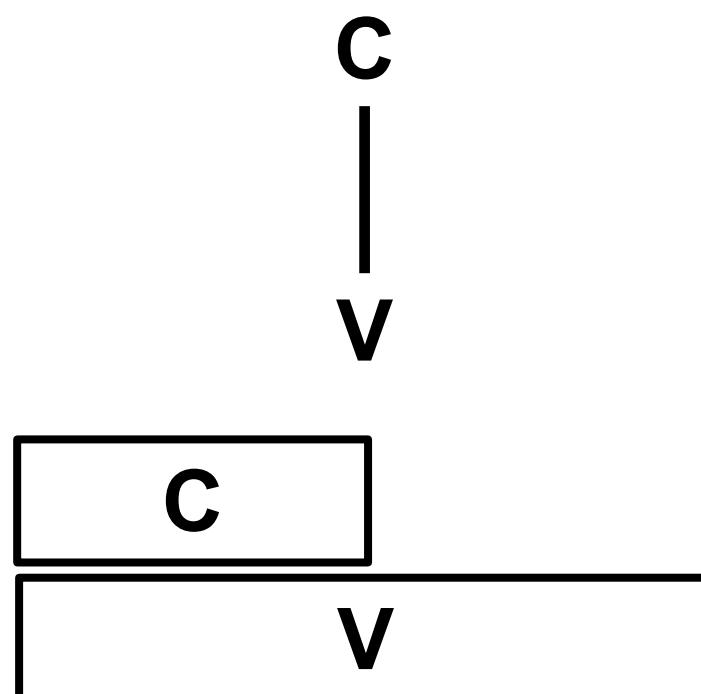
CV vs. VC syllables

in-phase

[pa]	
LIPS	Labial closure
TONGUE TIP	
TONGUE BODY	pharyngeal wide

anti-phase

[ap]	
LIPS	labial closure
TONGUE TIP	
TONGUE BODY	pharyngeal wide



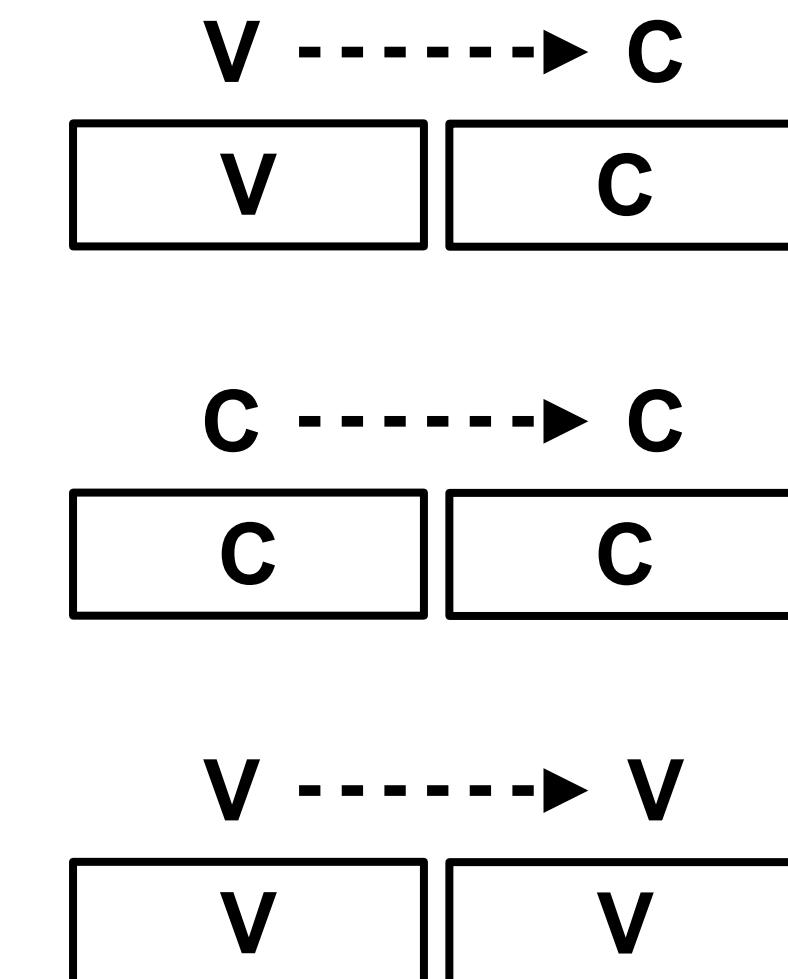
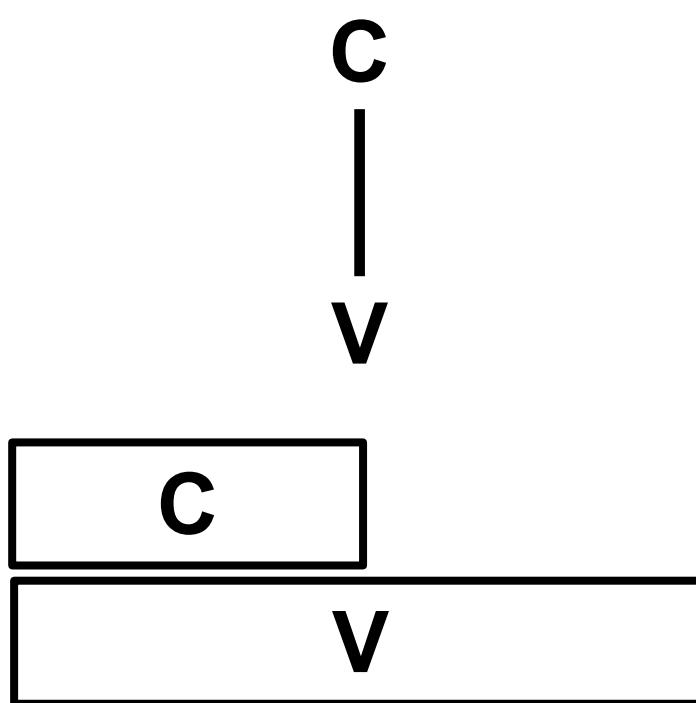
CV vs. VC syllables

in-phase

[pa]	
LIPS	Labial closure
TONGUE TIP	
TONGUE BODY	pharyngeal wide

anti-phase

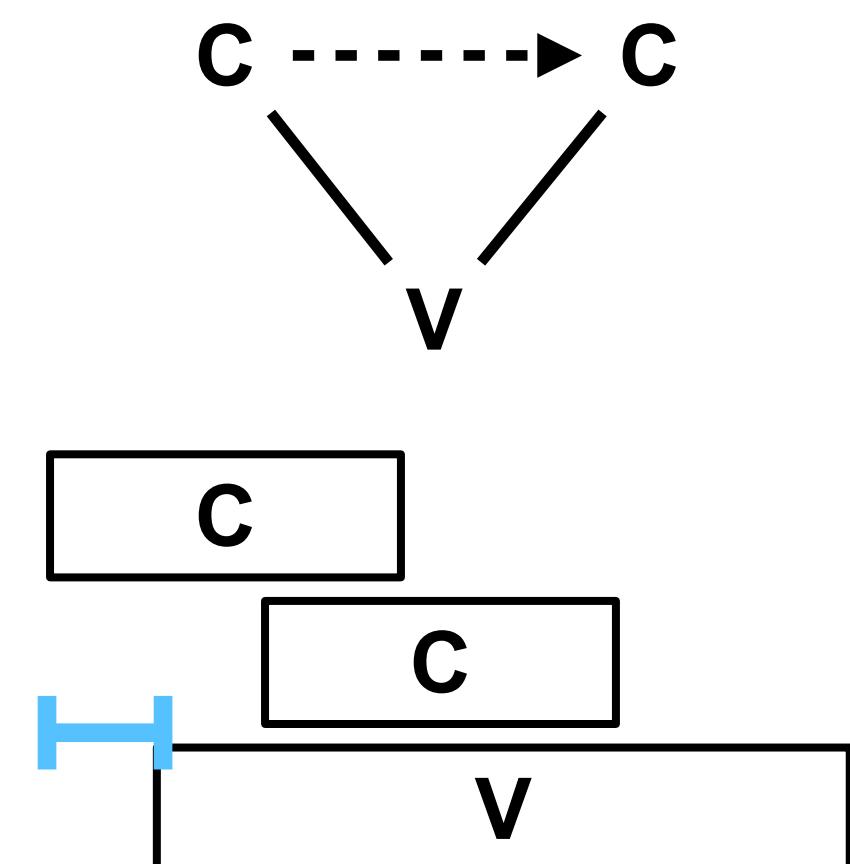
[ap]	
LIPS	labial closure
TONGUE TIP	
TONGUE BODY	pharyngeal wide



What about clusters?

- Empirically, onset clusters overlap

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



What about tone?

- Empirically, V lags following C
 - (In *lexical tone* languages only)

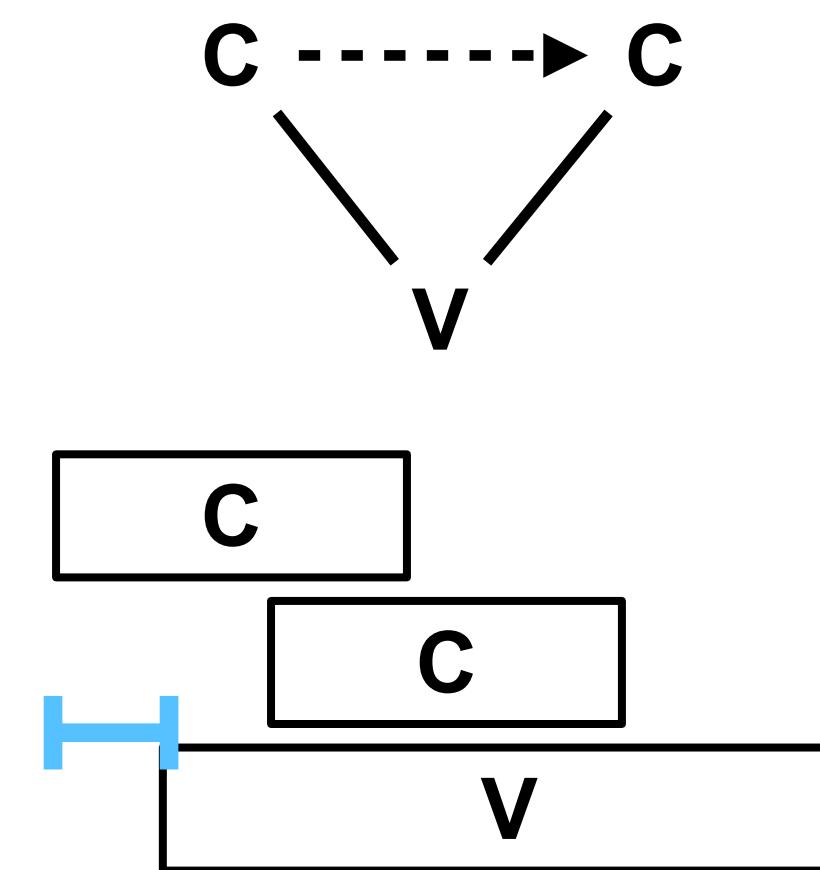
/pá/	
LIPS	labial closure
TONGUE TIP	
TONGUE BODY	pharyngeal wide
pitch (?)	high

Competitive coupling account

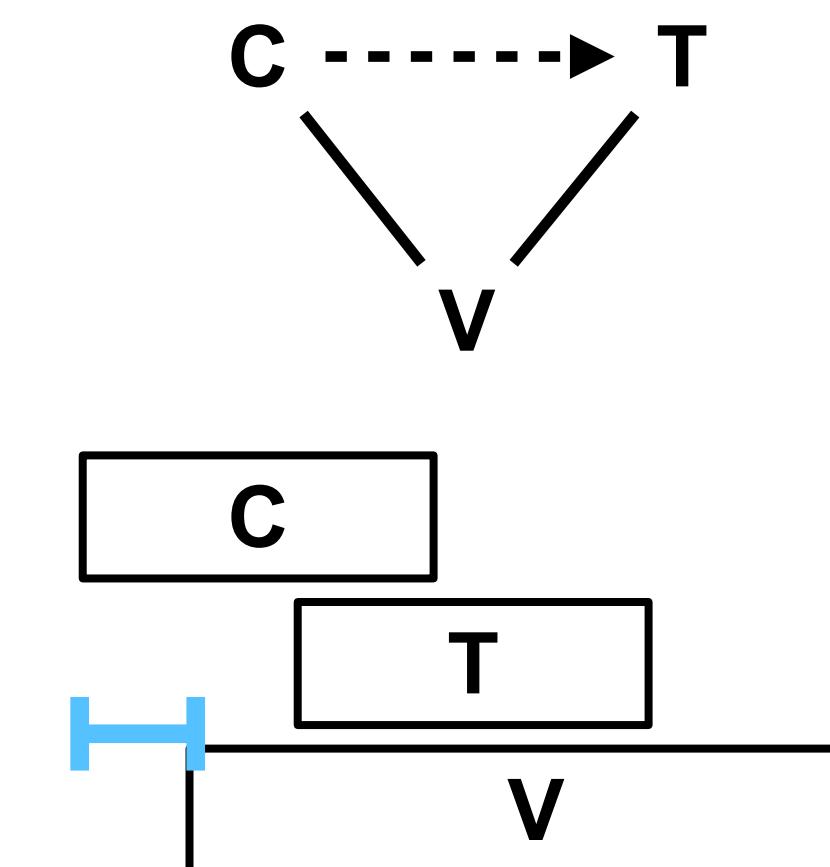


- Unifies clusters and tone (neat for typology)
- Unifies syllables (and up?), contrast, and planning

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

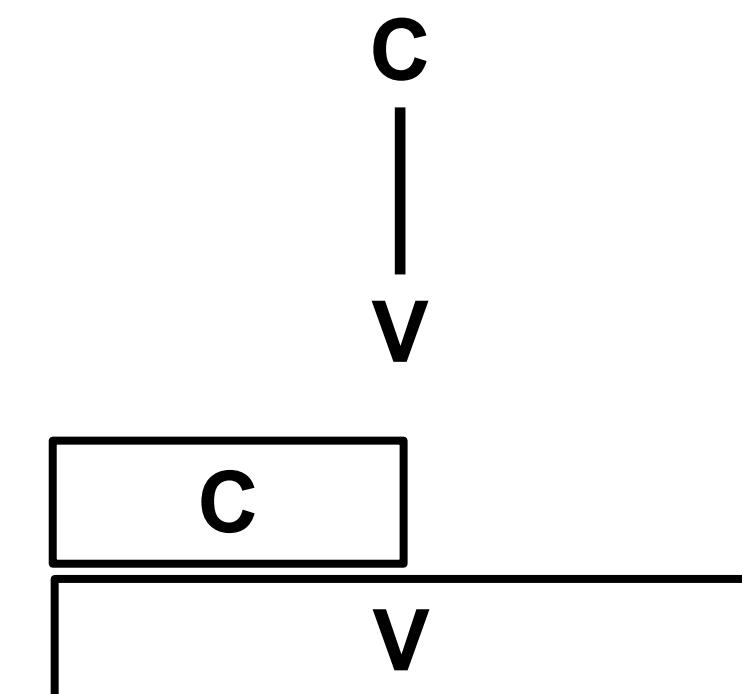
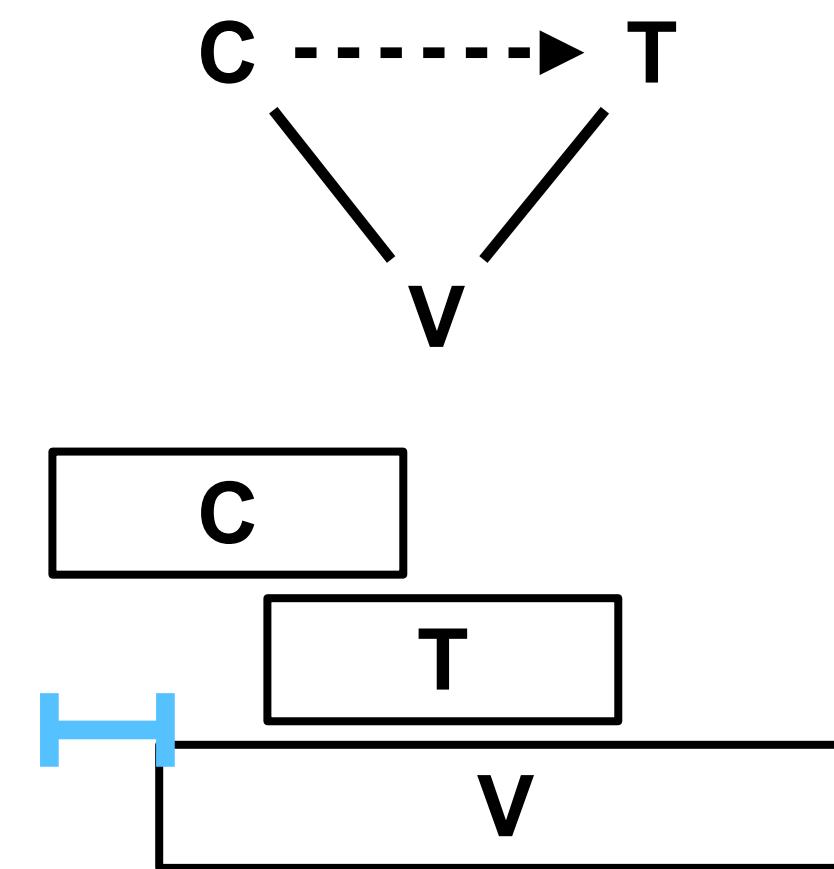


/pá/	
LIPS	labial closure
TONGUE TIP	
TONGUE BODY	pharyngeal wide
pitch (?)	high



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



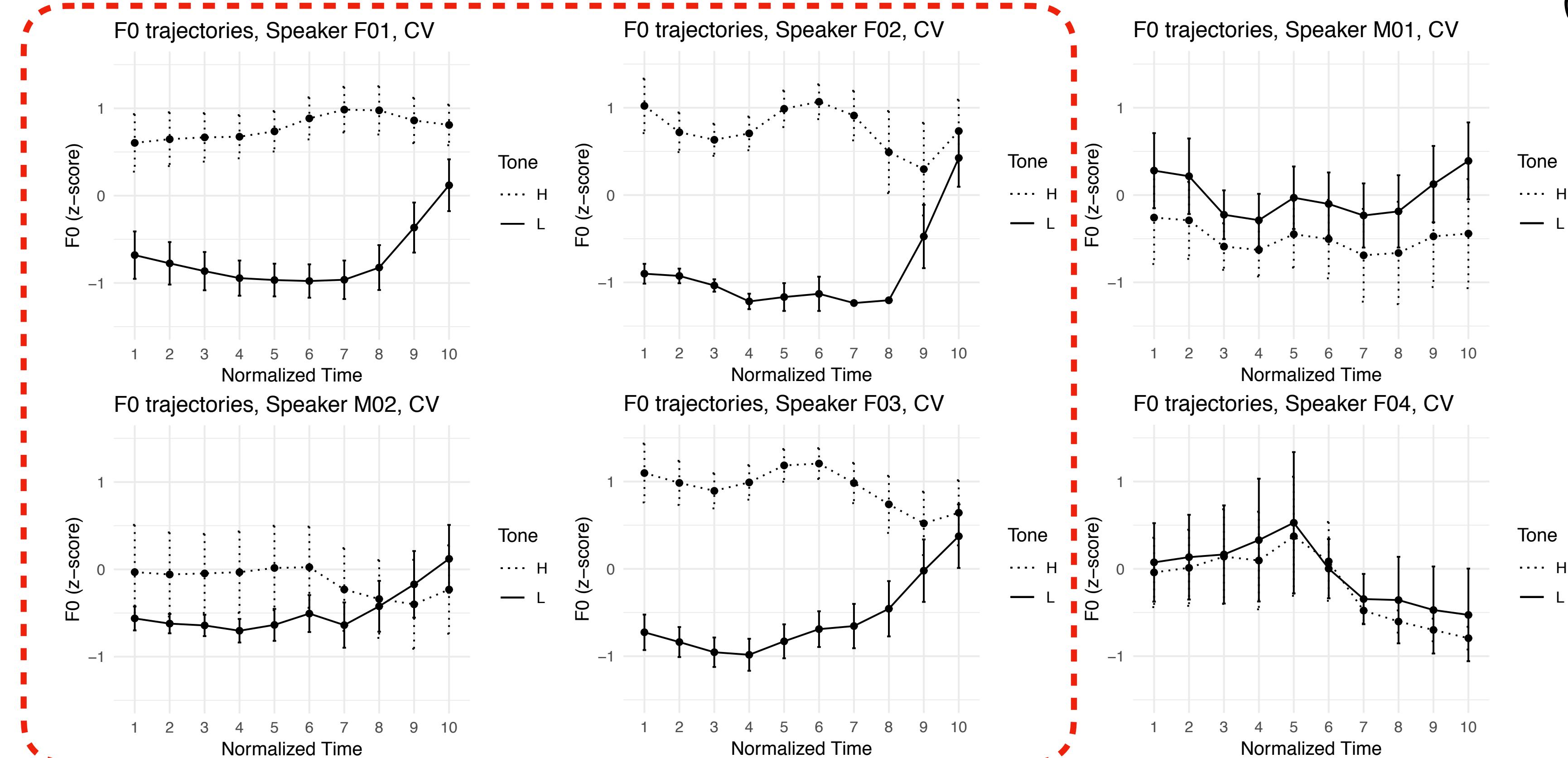
The perfect test case?

A language where some speakers produce tone and others don't

(Geissler 2019, 2021)

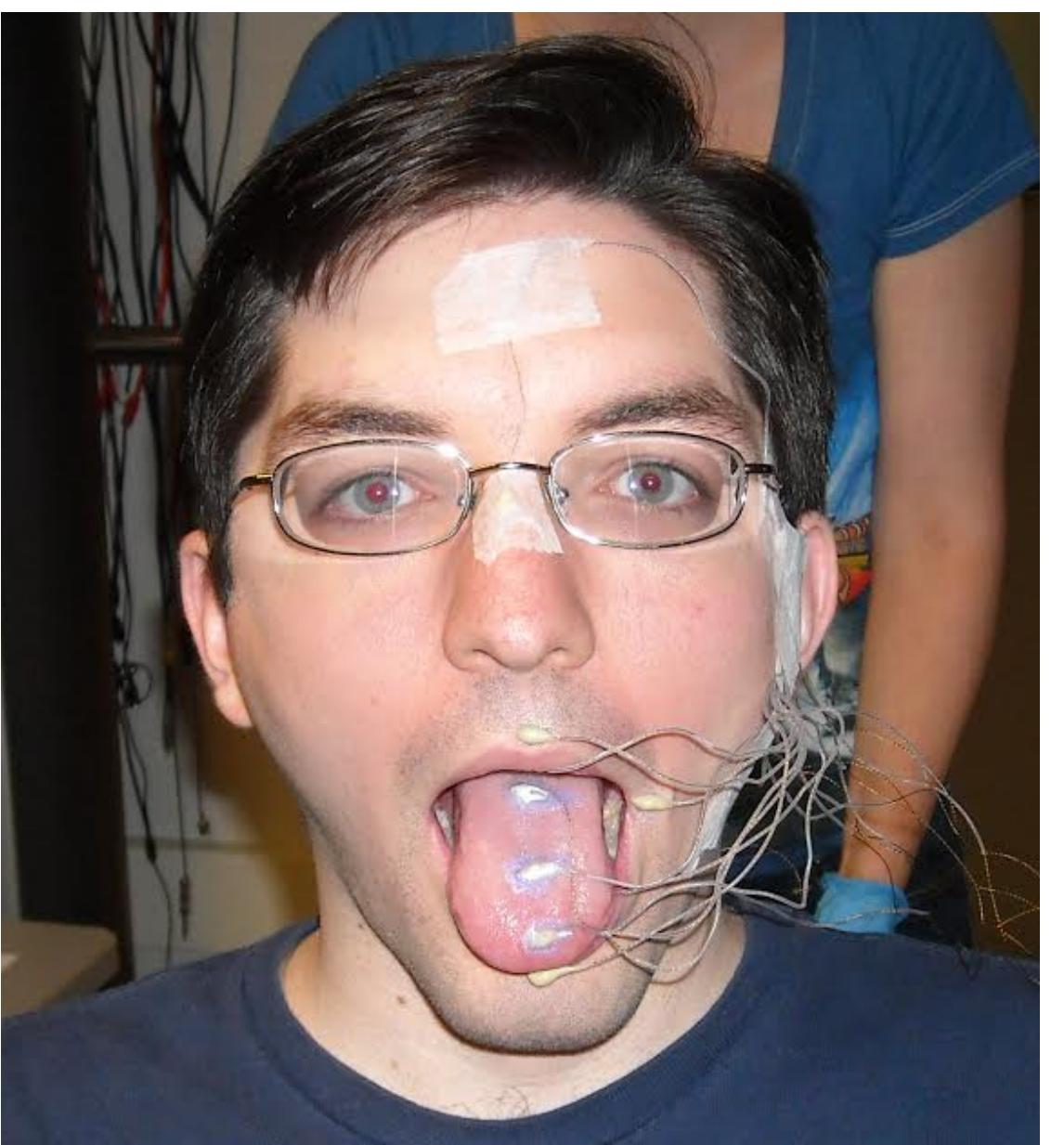
- 4 speakers produce a tone contrast, two do not (images: /mV/)

(Geissler et al. 2021)



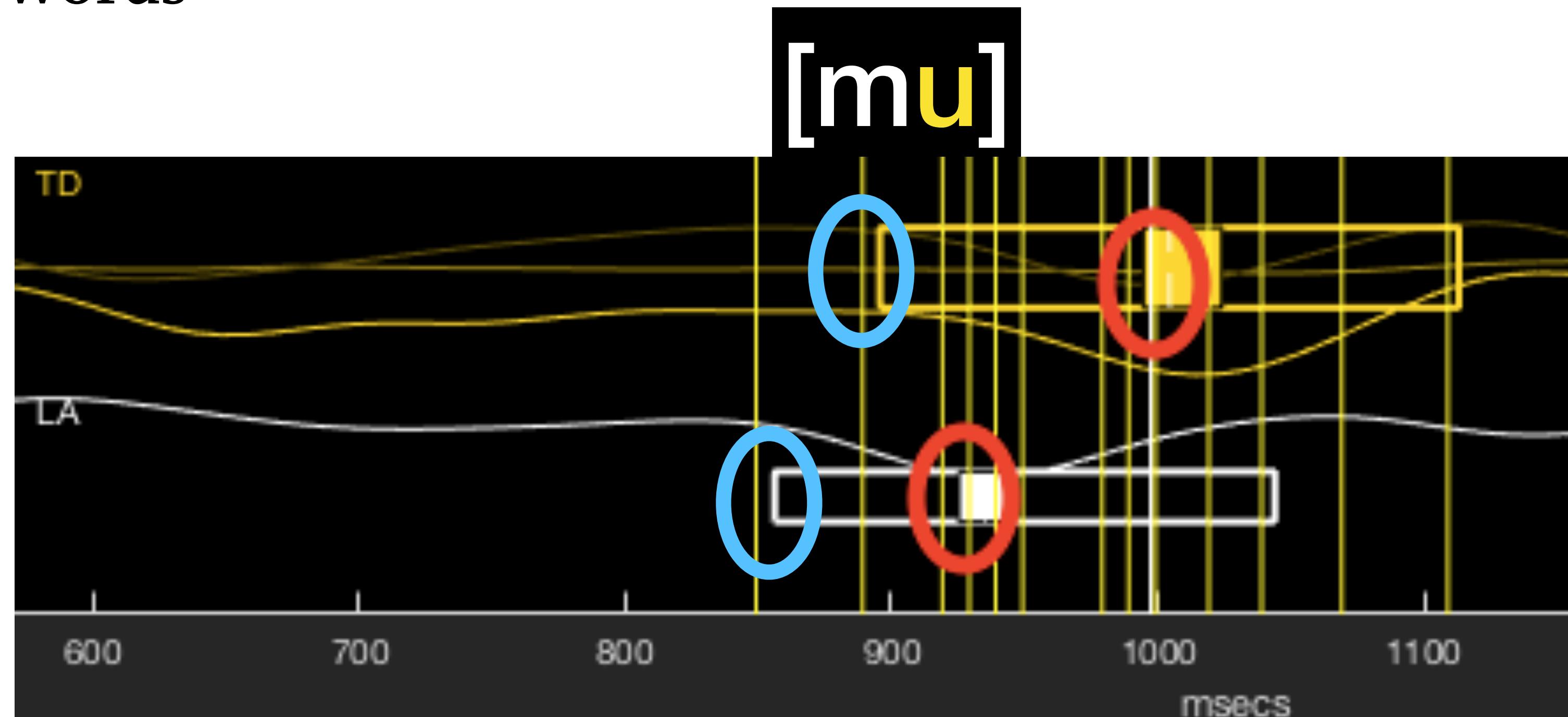
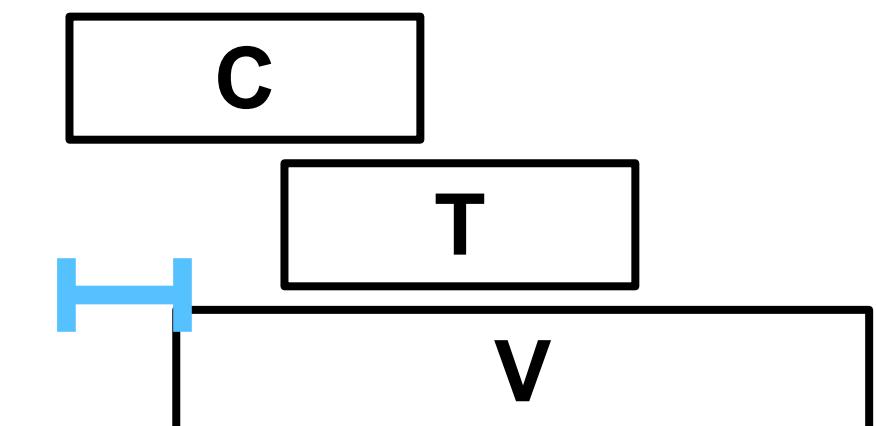
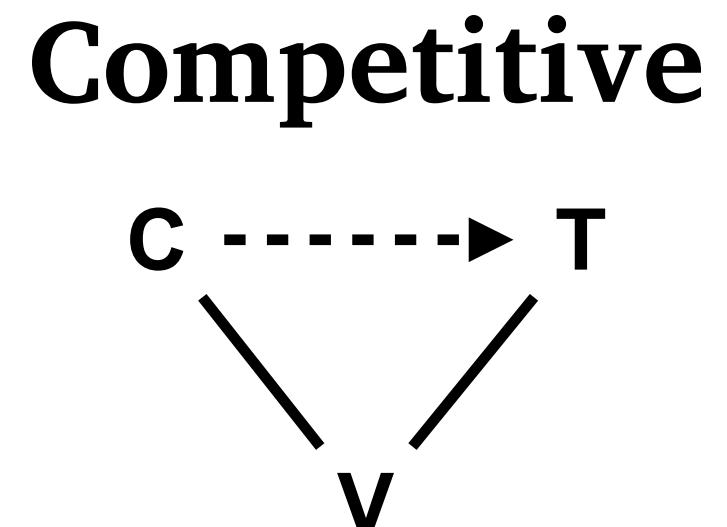
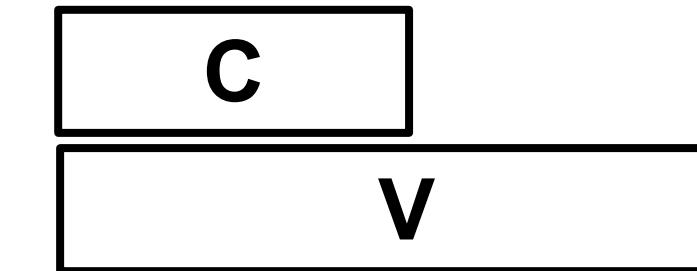
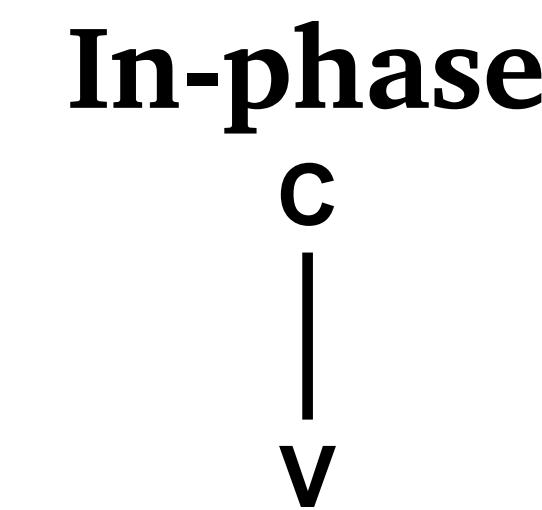
EMA study articulatory trajectories

- [p p^h m]: distance between lip sensors
- [i]→[u o a]: tongue dorsum retraction
- H, L tones; 1- and 2-syllable words
- C-V lag as diagnostic of tone



Tongue Dorsum
front
↓
back

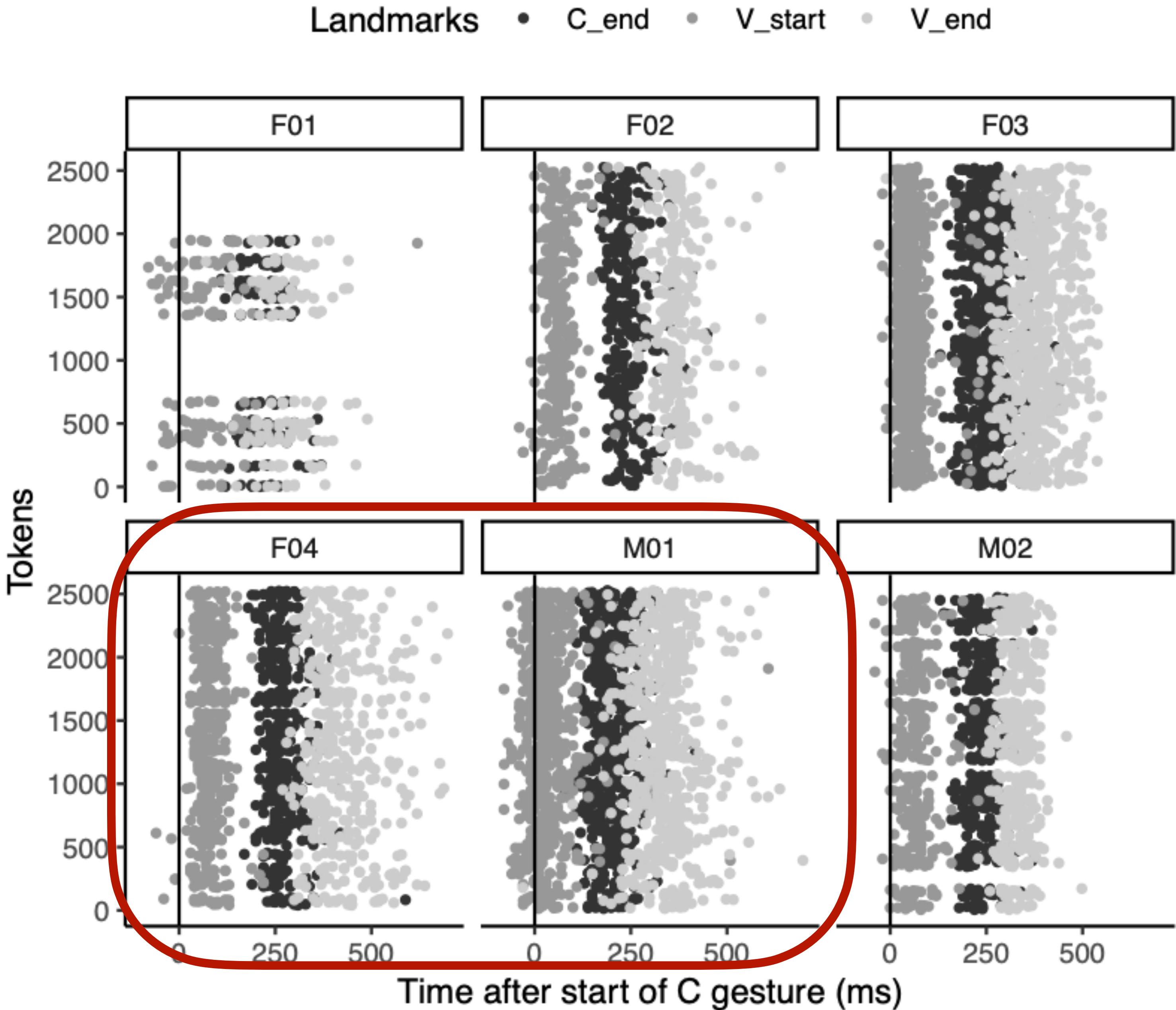
Lip Aperture
open
↓
closed



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

Results: C-V lag

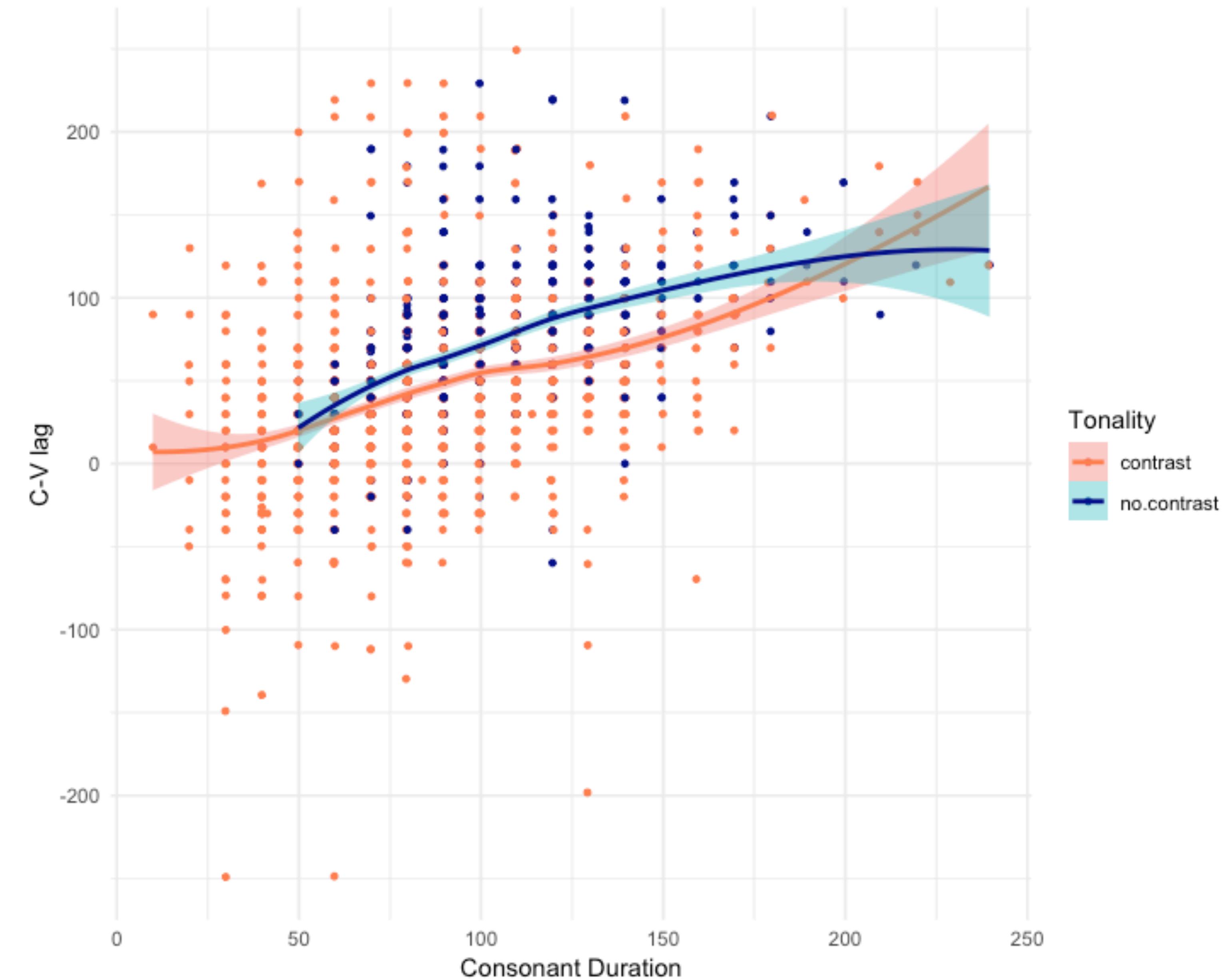
- There is a positive C-V lag... for speakers with *and* without the tone contrast (and in both tones)
- Competitive Coupling has no explanation for the 50ms lag



Effect of C duration on C-V lag

Results: C-V lag

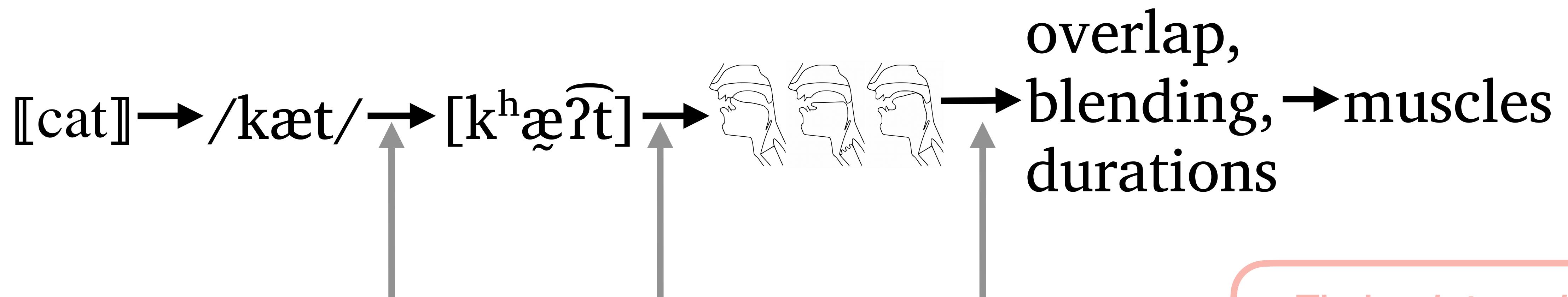
- C-V lag *does* increase with C duration
- so, the 50ms lag isn't just a fixed value
- intrinsic account: all speakers anti-phase (ish)
- extrinsic account: gestures and coordination both affect by speech rate



Roadmap

- Biography
- Phonology, phonetics, and time
- Types of evidence
 - Intergestural timing—Tibetan tonality
 - **Duration tradeoffs—Northern Saami**
 - Simulated trajectories
- Conclusion

Representational units



Symbolic Phonology	<i>Phonemes & Features</i>	" <i>Phonetic implementation</i> "
Articulatory Phonology	<i>Gestures & Timing Relations</i>	<i>Task Dynamics</i>
XT/3C	<i>Phonemes & Features</i>	<i>Seconds</i>

Timing *internal* or
external to phonology?

Both categorical and
continuous timing?

Which better fits
articulatory data?

Northern Sámi quantity distinctions

- 2 vowel lengths
- 3 (!!!) phonological consonant lengths:
 - Q1: [v̄iesu] ‘house (acc sg)’
 - Q2(~Q1): [v̄ies:u] ‘house (nom sg)’
 - Q2(~Q3): [r̄uo:s:a] ‘cross (acc sg)’
 - Q3: [r̄uos:a] ‘cross (nom sg)’
- Notice the [r̄uo]~[r̄uo:]?

Northern Sámi quantity distinctions

- 2 vowel lengths
- 3 (!!!) phonological consonant lengths:
 - Q1: [v̄iesu] ‘house (acc sg)’
 - Q2(~Q1): [v̄ies:u] ‘house (nom sg)’
 - Q2(~Q3): [r̄uo:s:a] ‘cross (acc sg)’
 - Q3: [r̄uos:a] ‘cross (nom sg)’
- Notice the [ūo]~[ūo:]? [nom sg] has a floating mora

Confirm phonetically

- Predict:

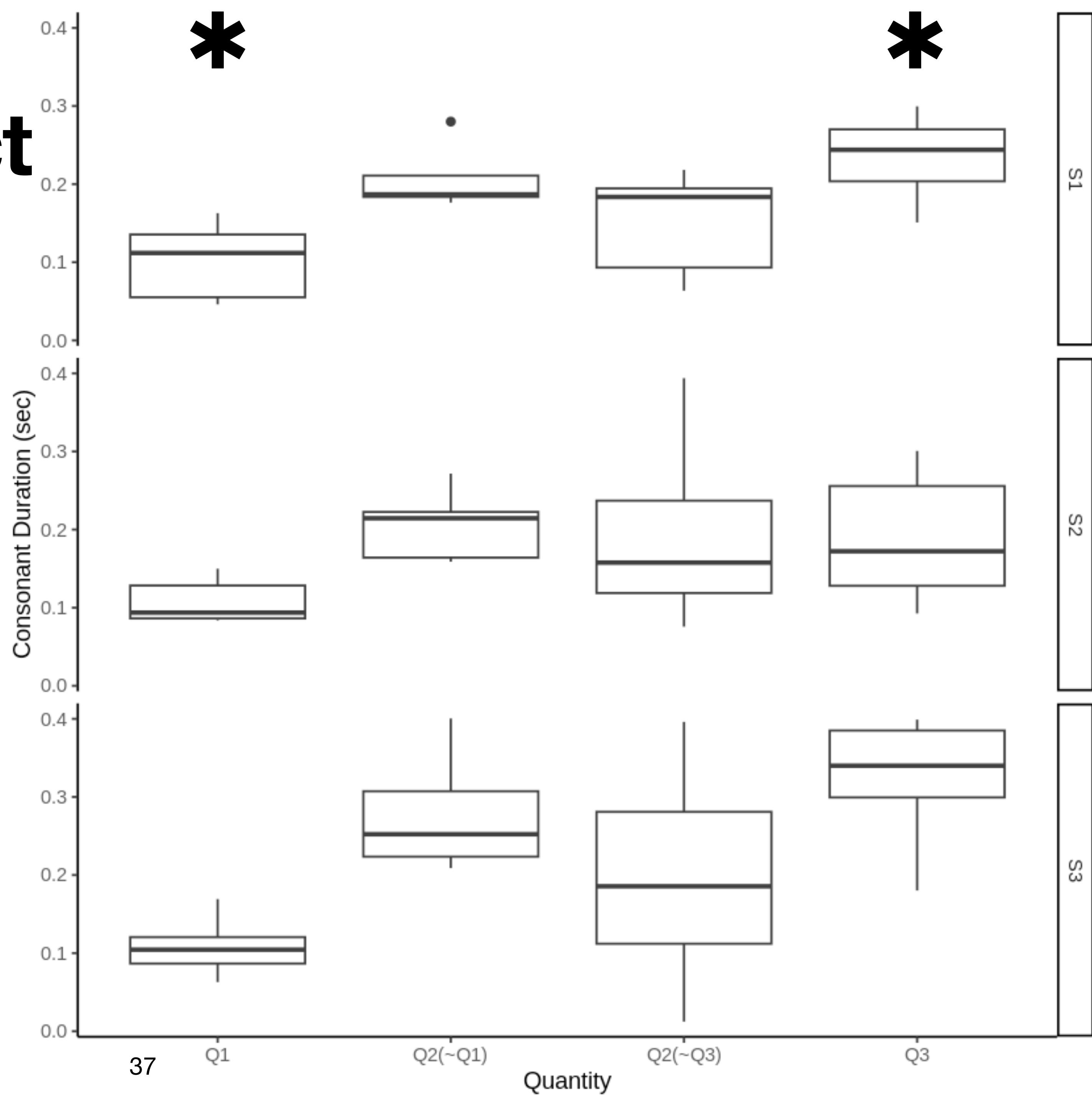
$Q_1 < Q_2(\sim Q_1) = Q_2(\sim Q_3) < Q_3$

shortest ————— longest

Phonological effect

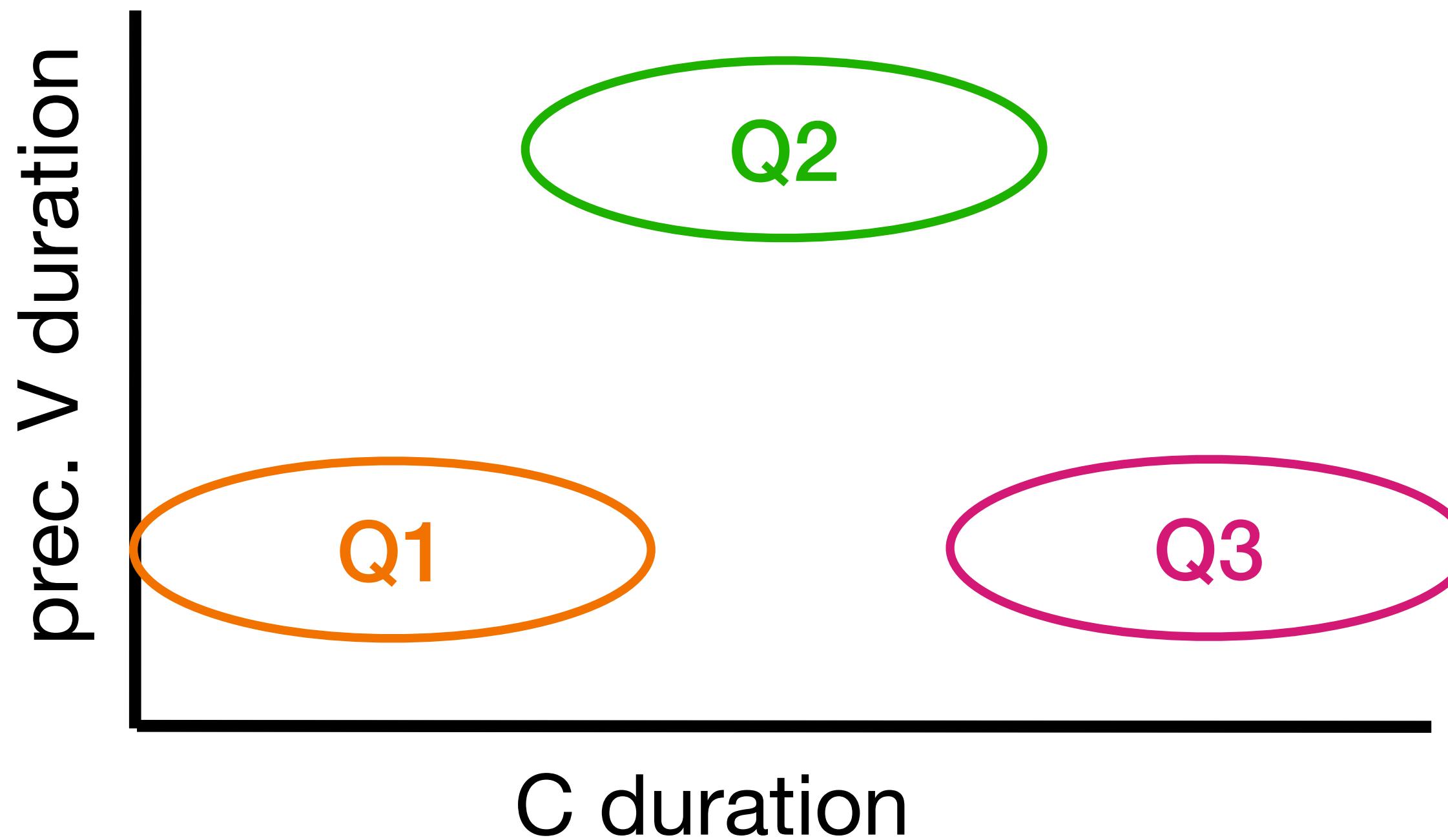
3 lengths? 3 lengths.

- Overall:
 - Q3 longer than Q2(\sim Q1) = Q2(\sim Q3)
 - longer than Q1
- S2 might have only two lengths; insufficient data

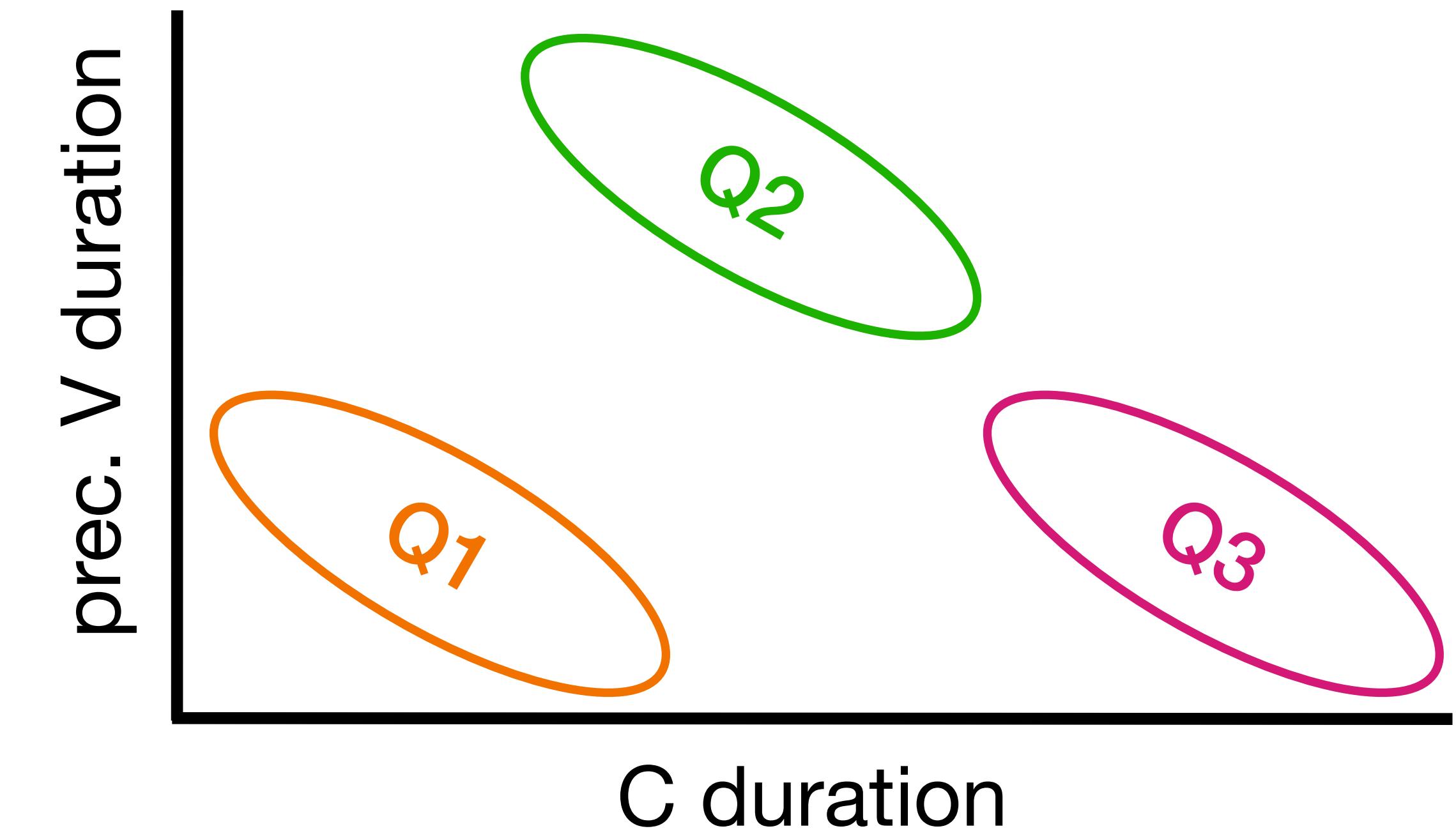


Phonological or phonetic?

- If phonological:



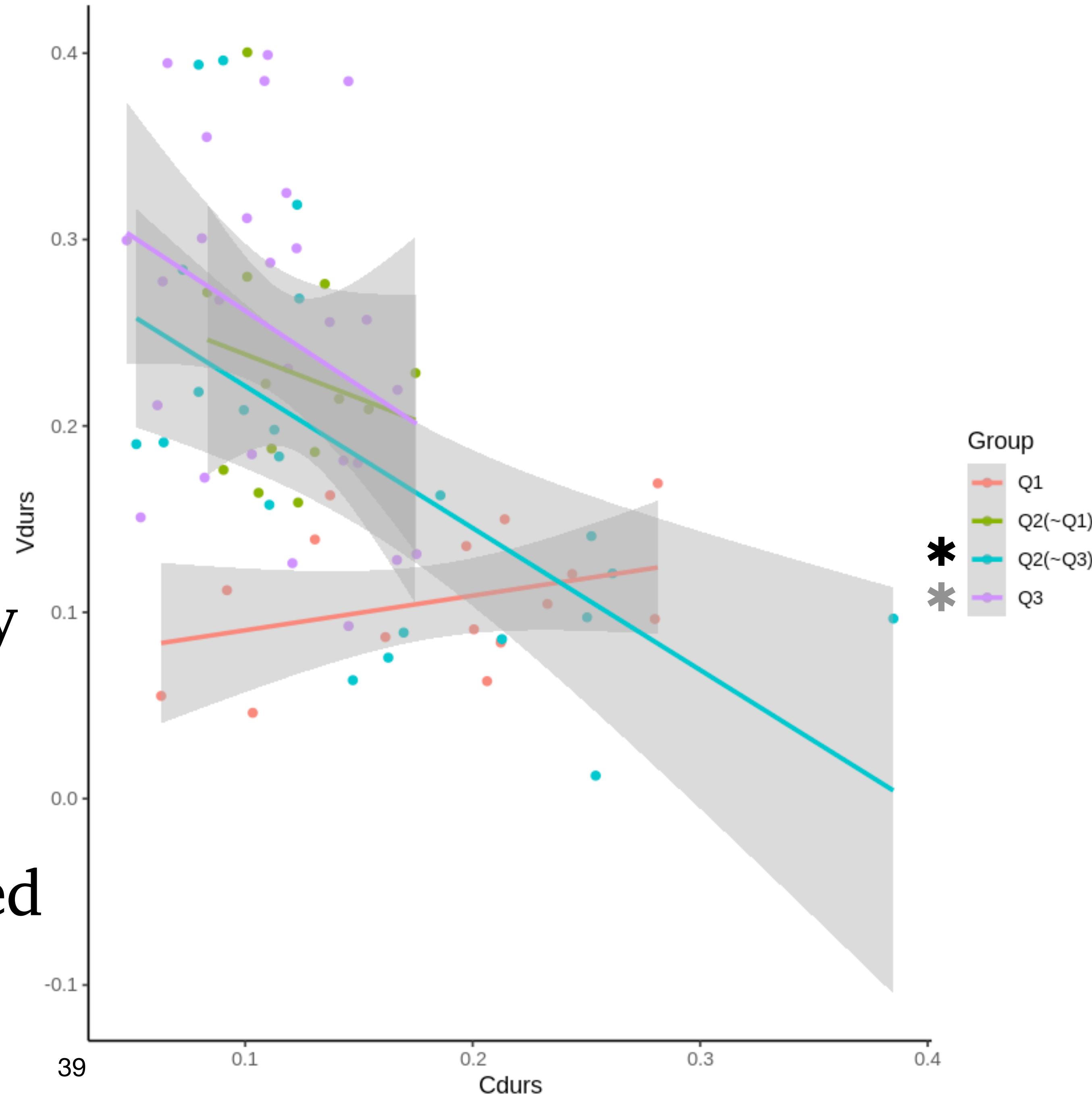
- If phonetic:



Phonetic effect

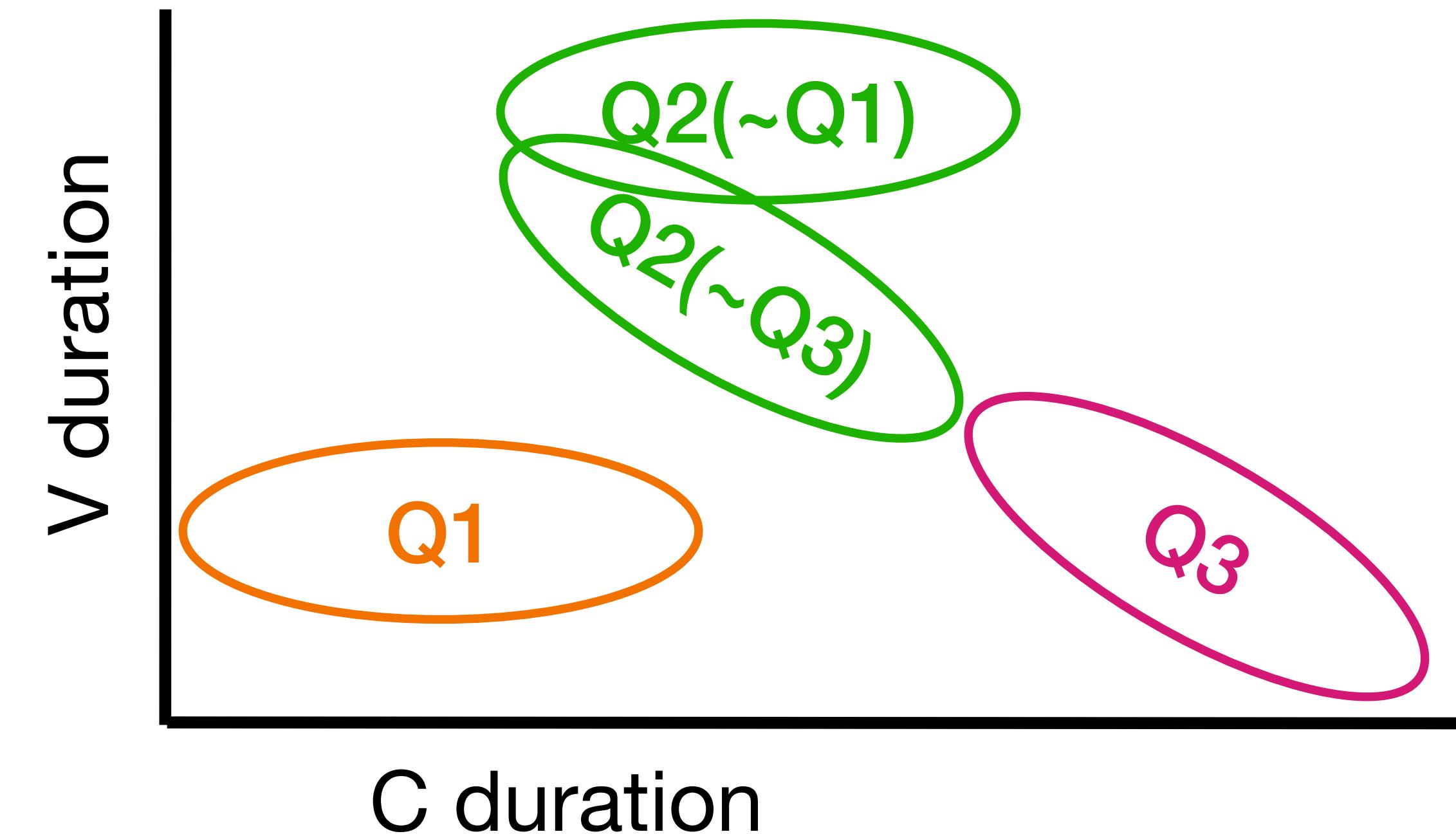
Inverse correlation

- Significant inverse relationships (V decrease when C increase) only in underlying Q3 Cs; driven by one speaker
- Trends in expected directions; more data needed



Phonological or phonetic?

- Phonological:
 - Q1
 - Q2(\sim Q1)
- Phonetic
 - Q3
 - Q2(\sim Q3)



Sámi summary

- For phonologically longest C's,
longer C's → shorter preceding V's
 - *this is over and above the phonological effect*
- For phonologically shortest C's,
no phonetic effect
 - *there is only the phonological effect*
- ... need more data...

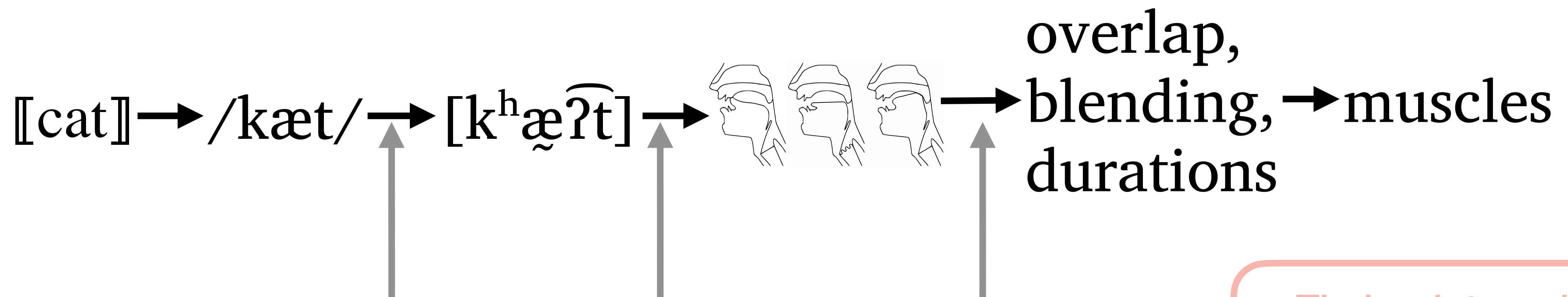
categorical and
continuous timing

only categorical timing

Roadmap

- Biography
- Phonology, phonetics, and time
- Types of evidence
 - Intergestural timing—Tibetan tonality
 - Duration tradeoffs—Northern Saami
 - **Simulated trajectories**
- Conclusion

Representational units



Symbolic Phonology	<i>Phonemes & Features</i>	<i>“Phonetic implementation”</i>
Articulatory Phonology	<i>Gestures & Timing Relations</i>	<i>Task Dynamics</i>
XT/3C	<i>Phonemes & Features</i>	<i>Seconds</i>

Timing *internal* or
external to phonology?

Both categorical and
continuous timing?

Which better fits
articulatory data?

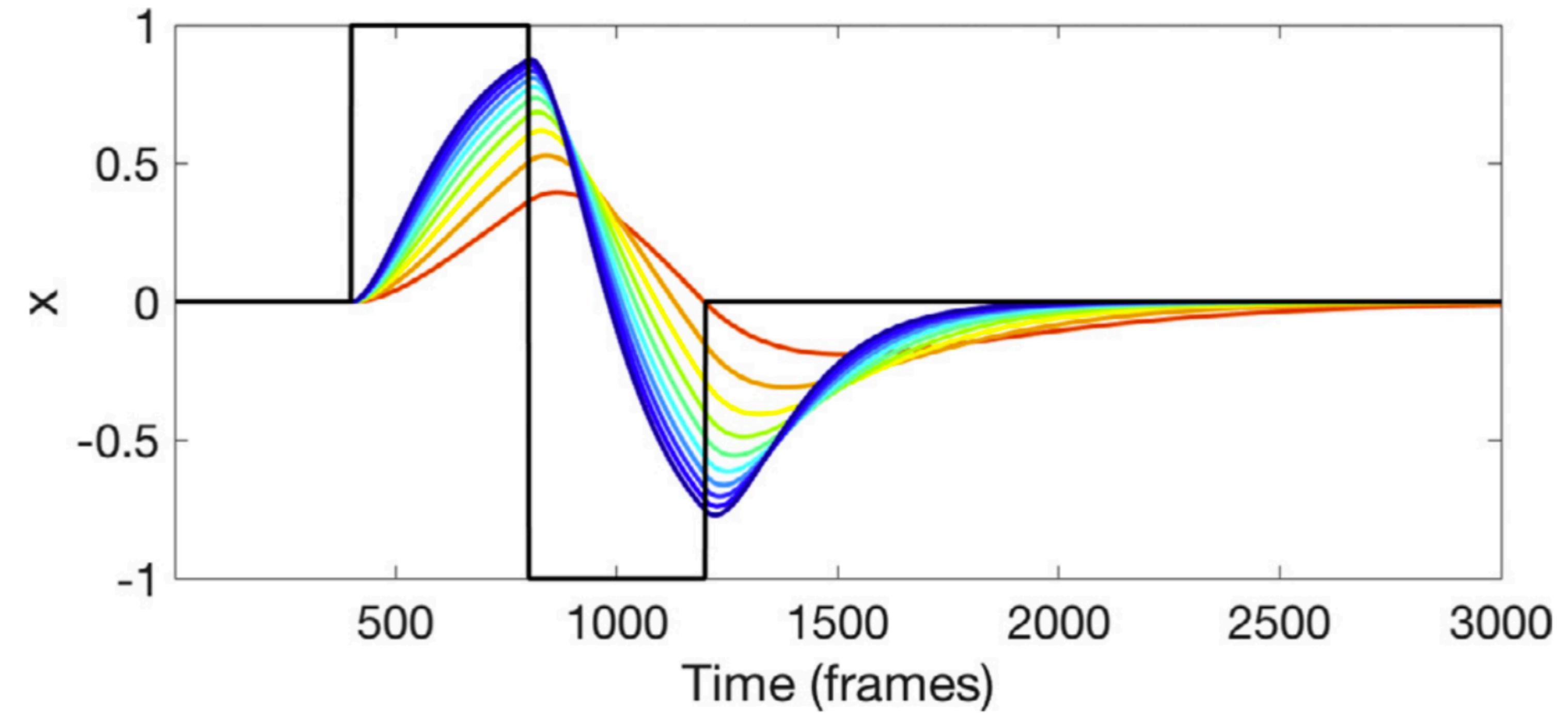
Task Dynamics (Articulatory Phonology)

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

- Model movement as critically-damped mass-spring oscillator
- Timing is *internal to the gesture* (sine waves are circles)

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

stiffness →
target →
velocity →
acceleration →



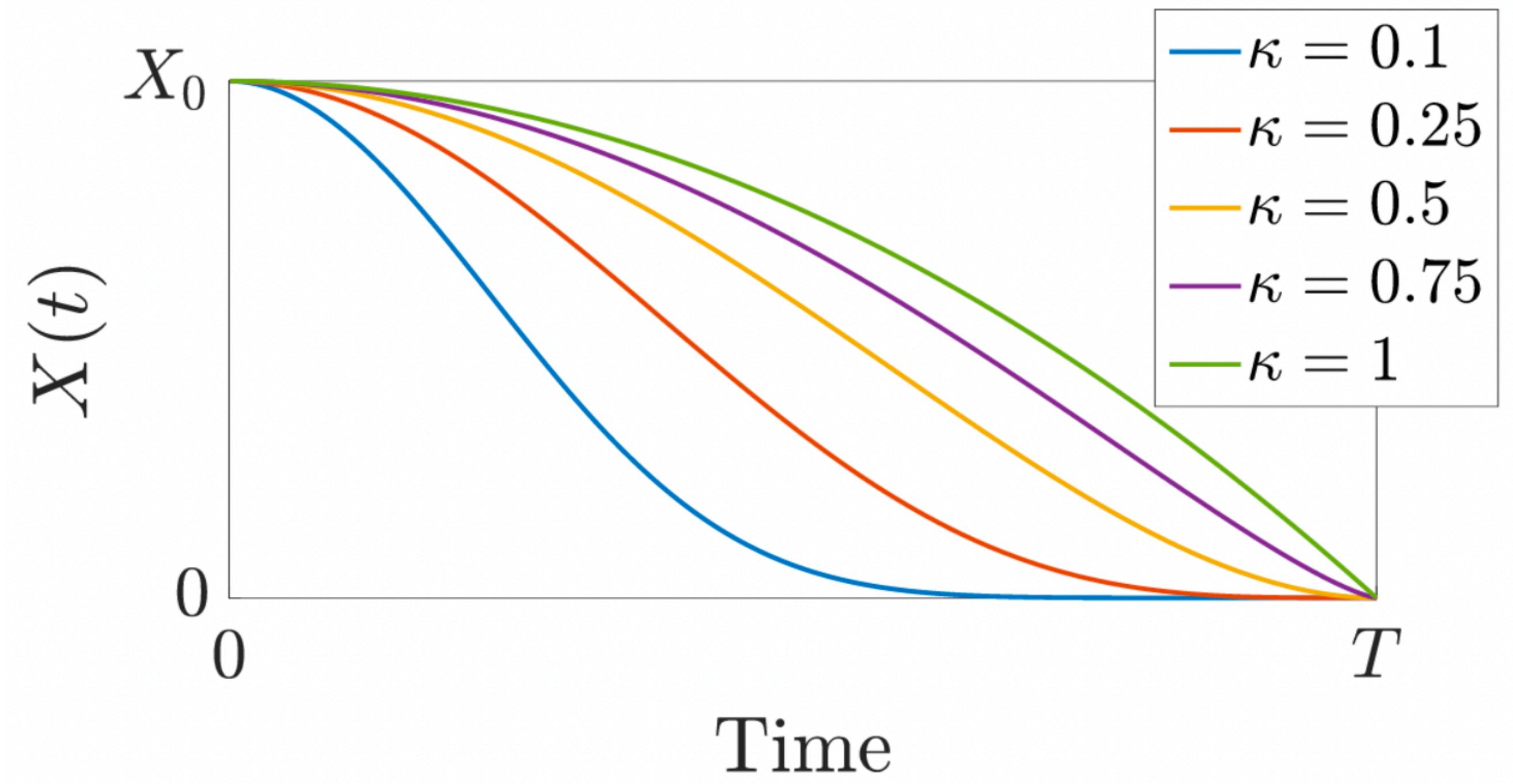
General Tau Theory (XT/3C)

(Lee & Reddish 1981, Turk & Shattuck-Hufnagel 2020)

- Model kinematics as gap-closing function
- Time only in regular, system-external time

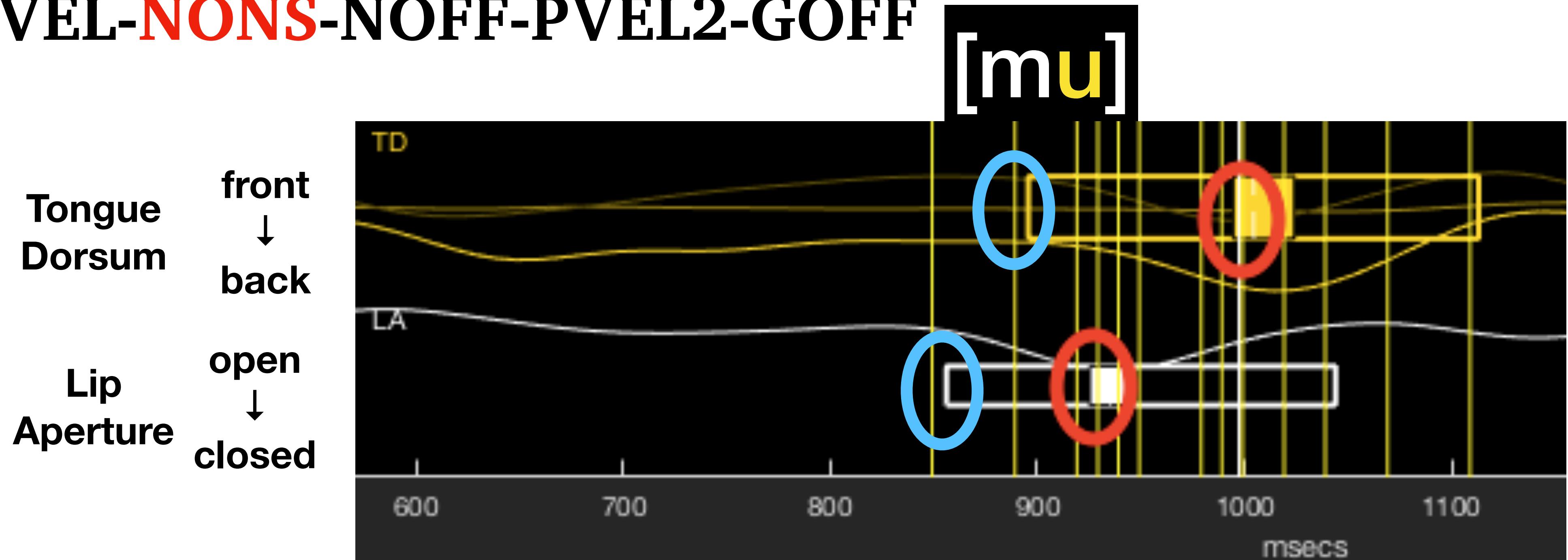
$$X(t) = X_0 \left(1 - \frac{t^2}{T^2} \right)^{\frac{1}{\kappa}}$$

Start position → position
end time → time
constant → κ



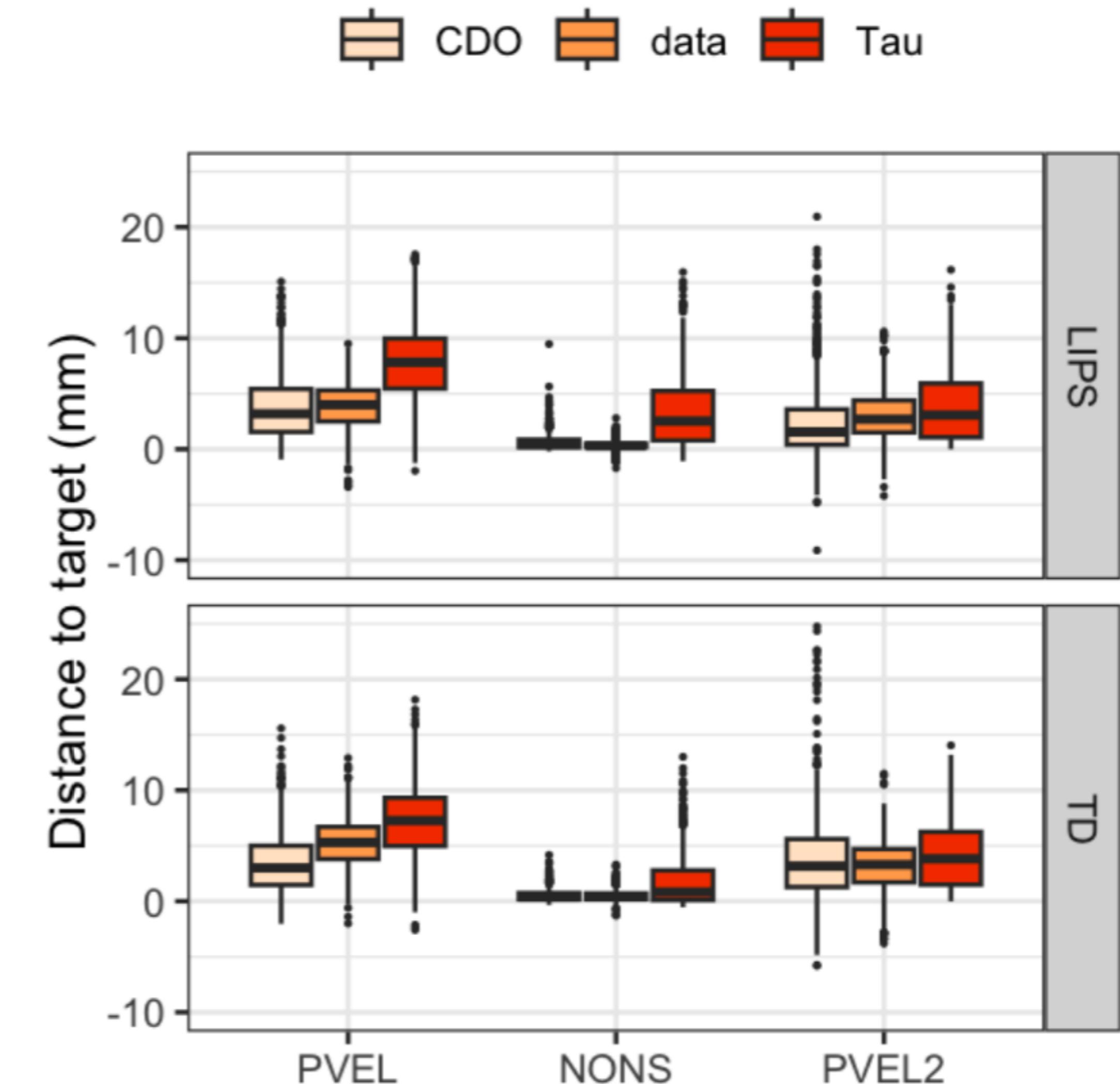
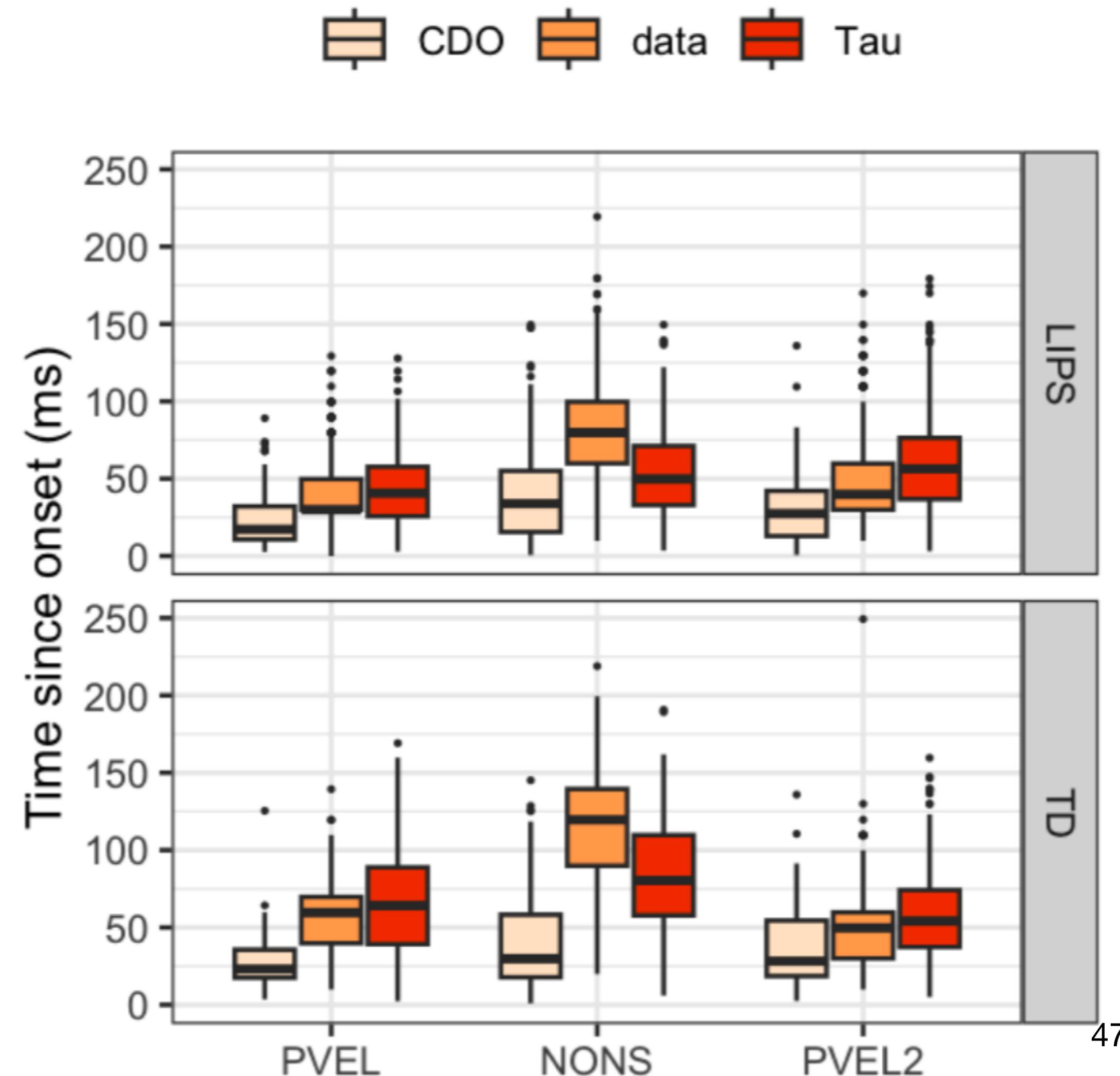
Which fits data better?

- Predicting landmarks from other landmarks:
 - **GONS-PVEL-NONS-NOFF-PVEL2-GOFF**



Which fits data better?

Time? Tau. Position? Oscillator?



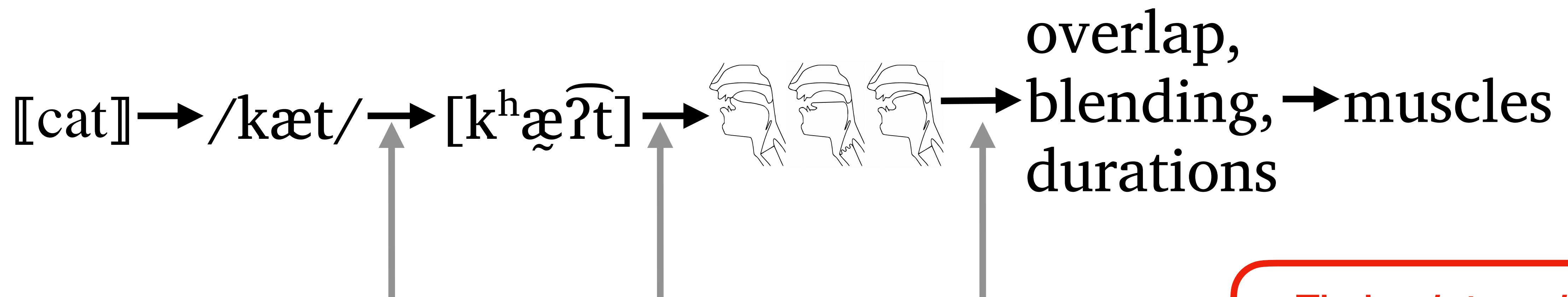
TD/Tau: Conclusion

- Work in progress!
 - Current: full trajectories, not just landmarks
- Results mixed
 - Tau better at *when* landmarks take place
 - TD better at *where* landmarks take place
 - This is weird

Roadmap

- Biography
- Phonology, phonetics, and time
- Types of evidence
 - Intergestural timing—Tibetan tonality
 - Duration tradeoffs—Northern Saami
 - Simulated trajectories
- Conclusion

Representational units



Symbolic Phonology	<i>Phonemes & Features</i>	<i>“Phonetic implementation”</i>
Articulatory Phonology	<i>Gestures & Timing Relations</i>	<i>Task Dynamics</i>
XT/3C	<i>Phonemes & Features</i>	<i>Seconds</i>



Timing *internal* or
external to phonology?



Both categorical and
continuous timing?



Which better fits
articulatory data?

សូមសម្រេច

Thank you!

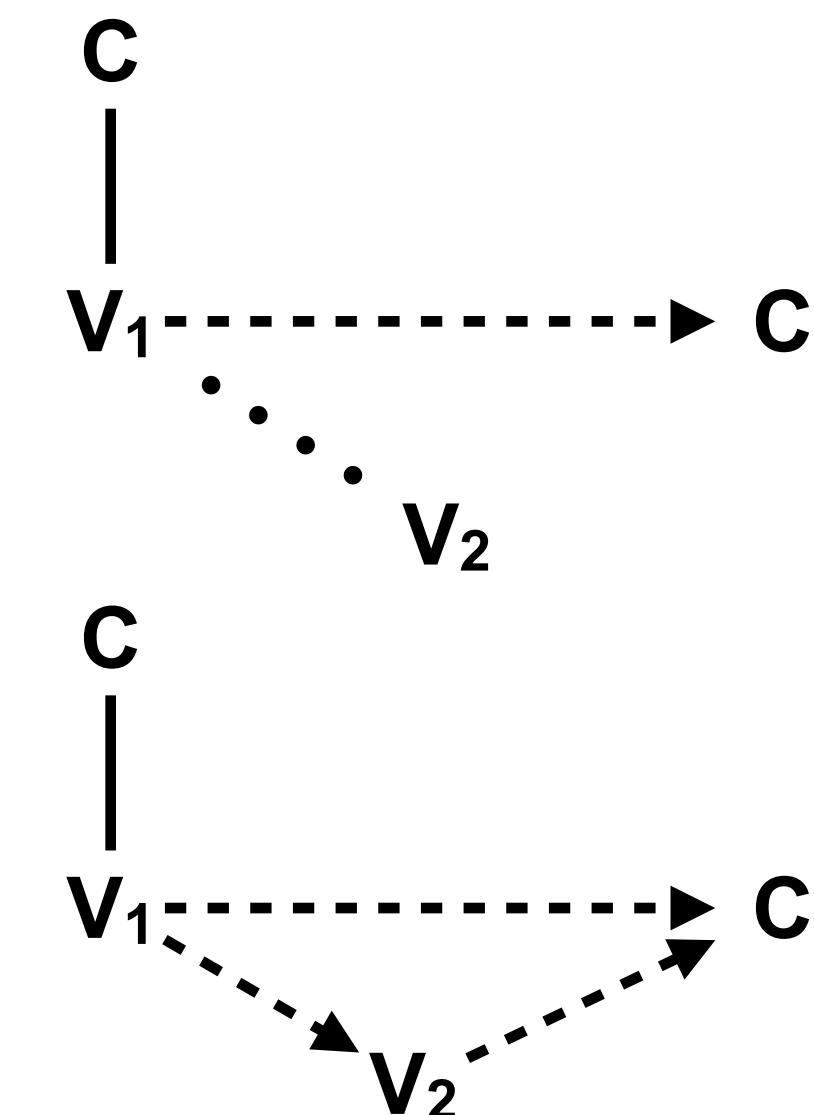
Pocket slides

What about diphthongs?

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

<five> /faɪv/

LIPS	labiodent. critical	labiodent. critical
TONGUE TIP		
TONGUE BODY	pharyngeal wide	palatal narrow
VELUM		
GLOTTIS	wide	



Articulatory study

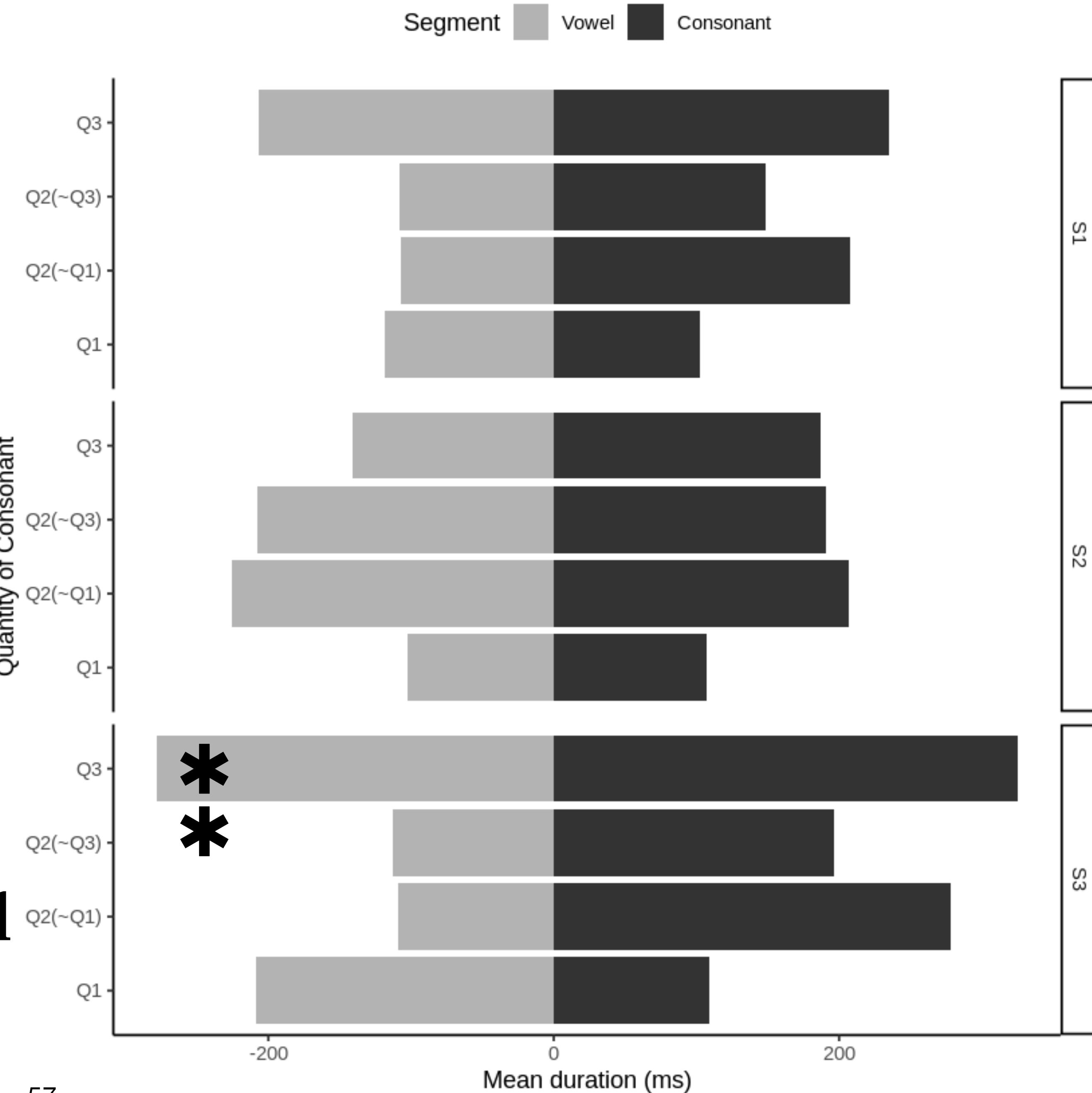
Geissler et al. (2021), Geissler (2021ch4)

- 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.

- Significant inverse relationships (V decrease when C increase) only in underlying Q3 Cs; driven by one speaker
- Trends in expected directions; more data needed



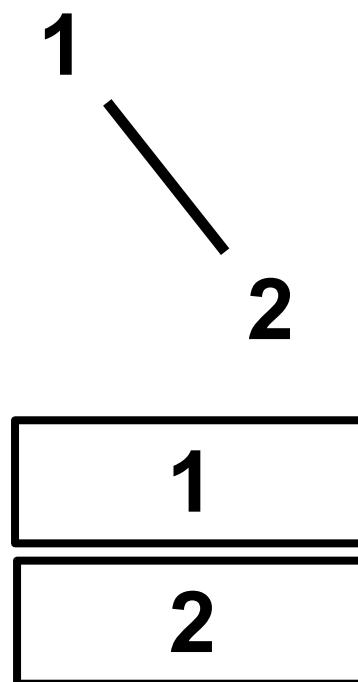
The temporal basis of complex segments

Shaw et al. 2019

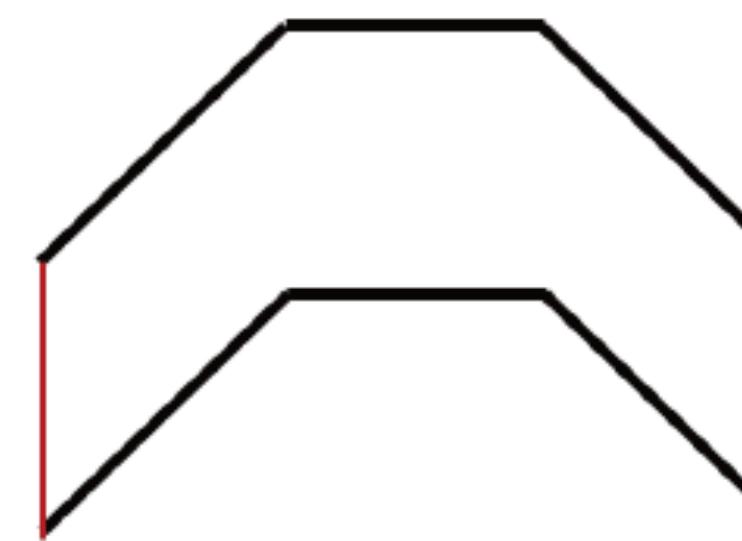
The temporal basis of complex segments

Shaw (2019): predictions

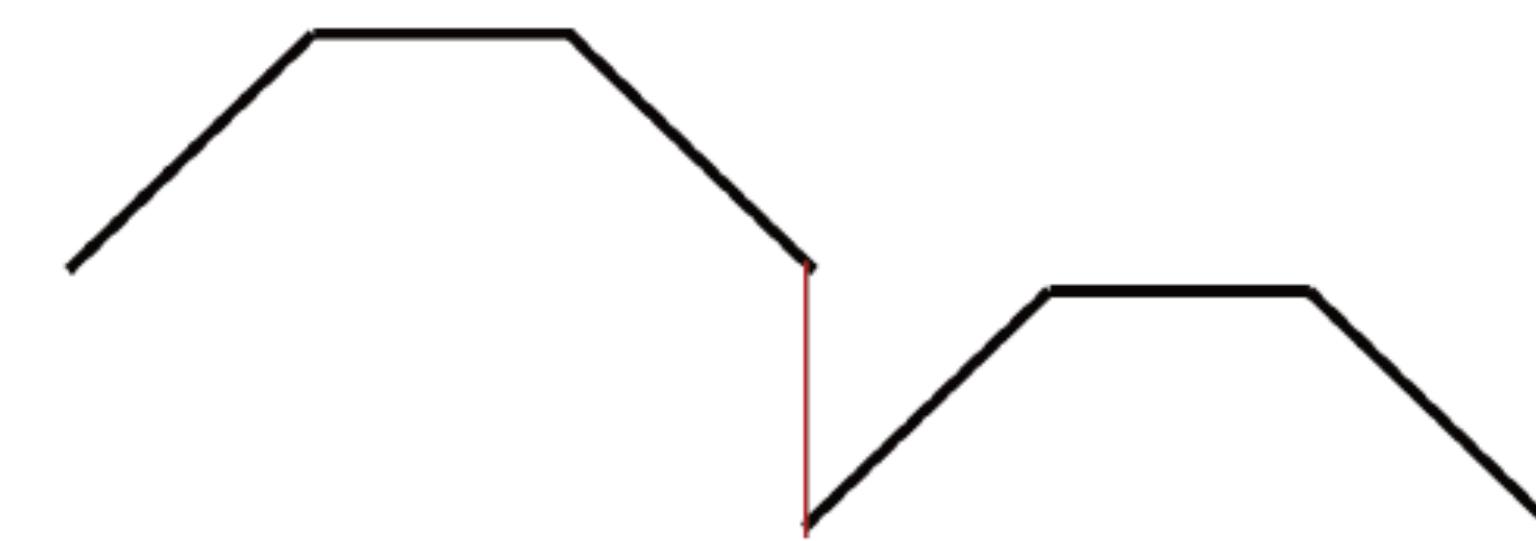
In-phase



(a) Complex segment—no lag



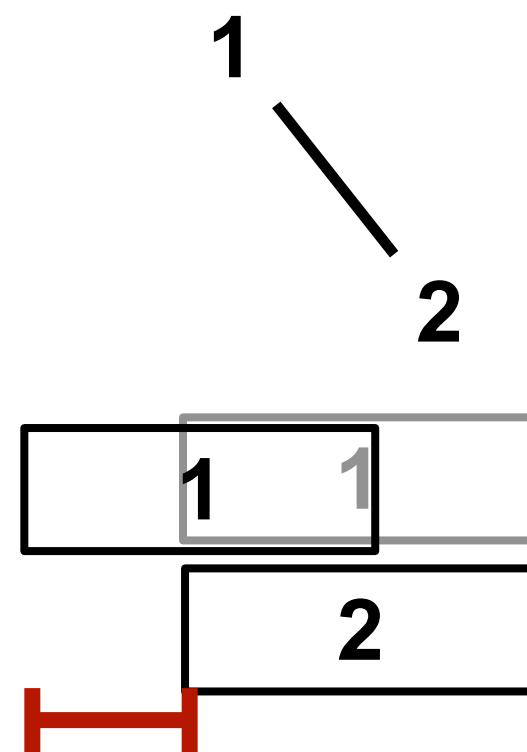
(b) Segment sequence—no lag



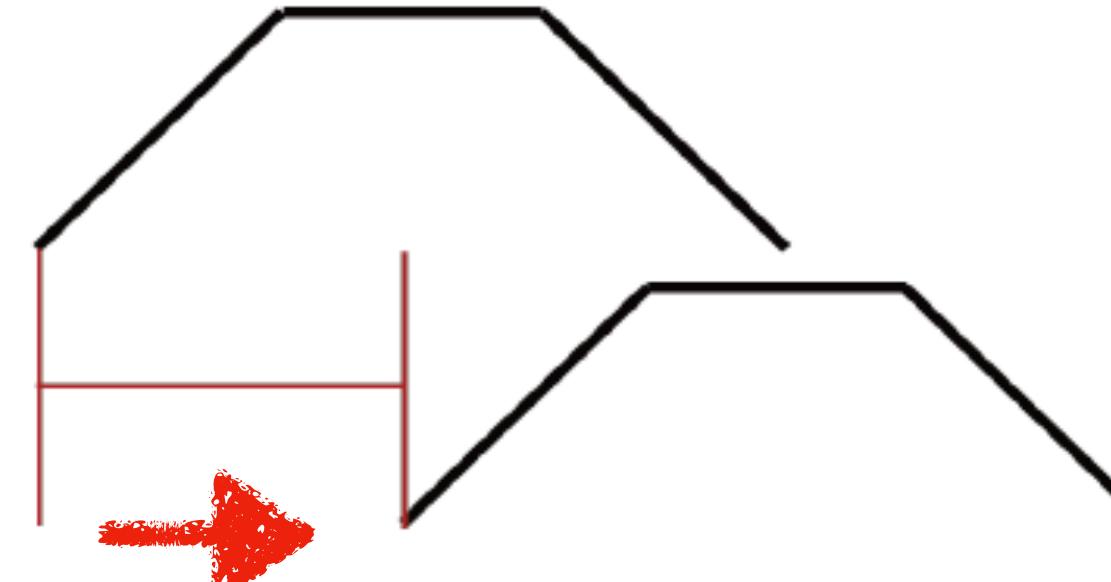
Anti-Phase



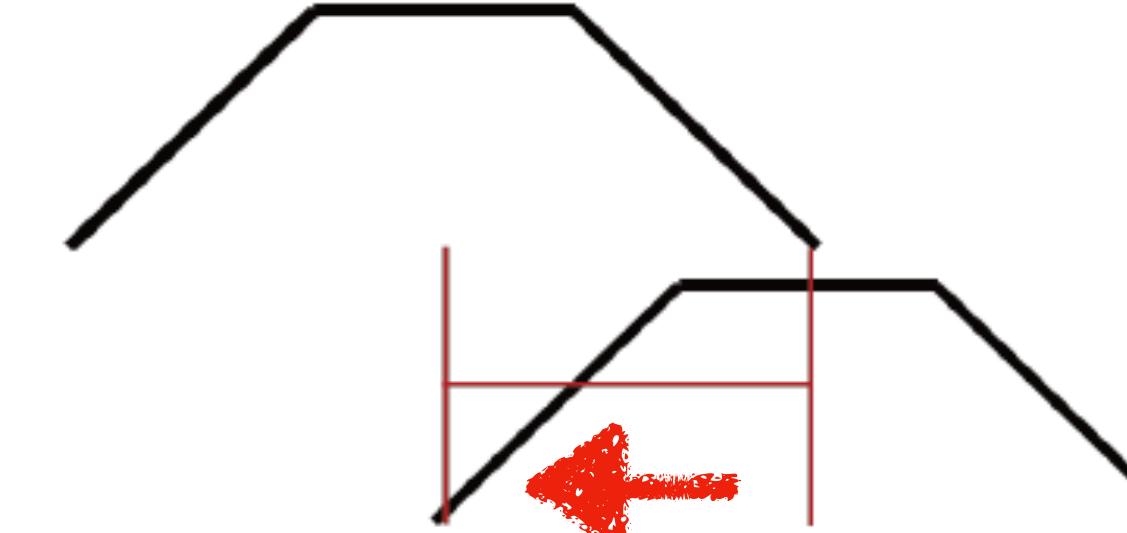
In-phase + lag
(offset)



(c) Complex segment—positive lag



(d) Segment sequence—negative lag



Anti-Phase - lag
(offset)

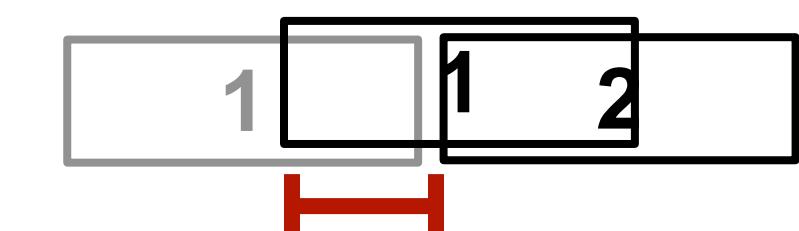


Figure 1: Hypothesized gestural coordination patterns for complex segments (a), (c) and segment sequences (b), (d)

The temporal basis of complex segments

Shaw (2019): results

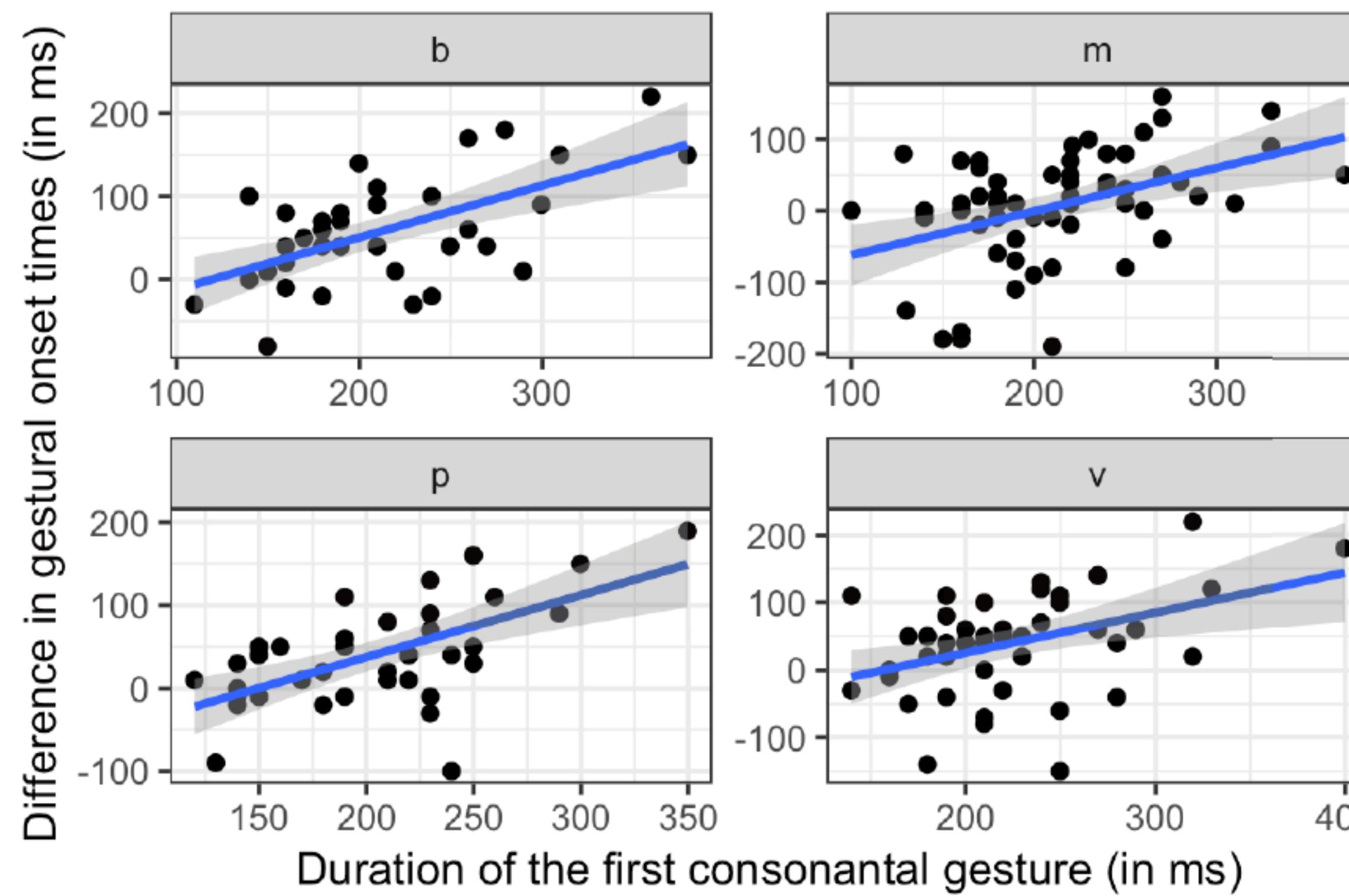


Figure 4: Correlations for the data from the English experiment

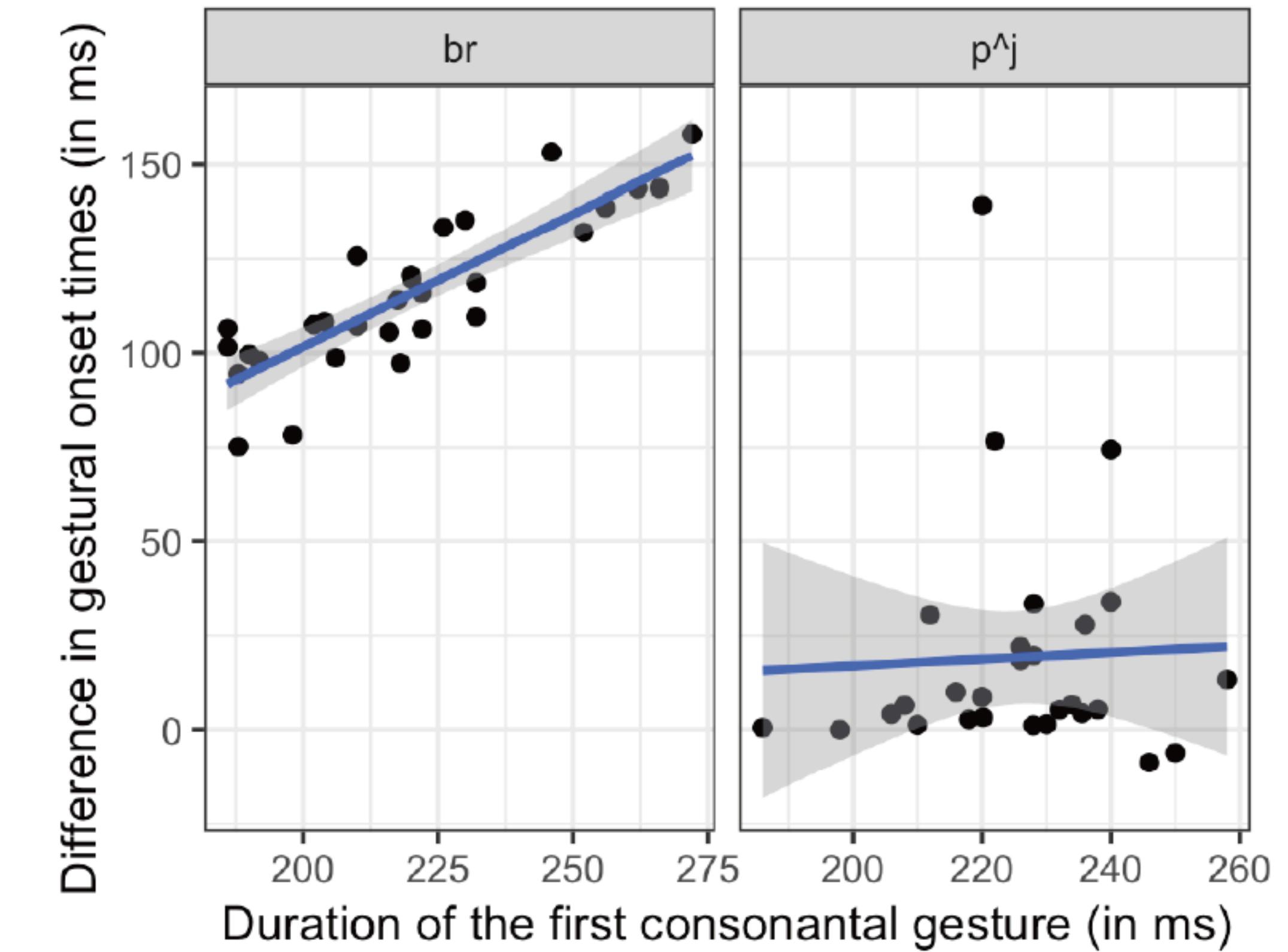


Figure 2: Correlations for the Russian data

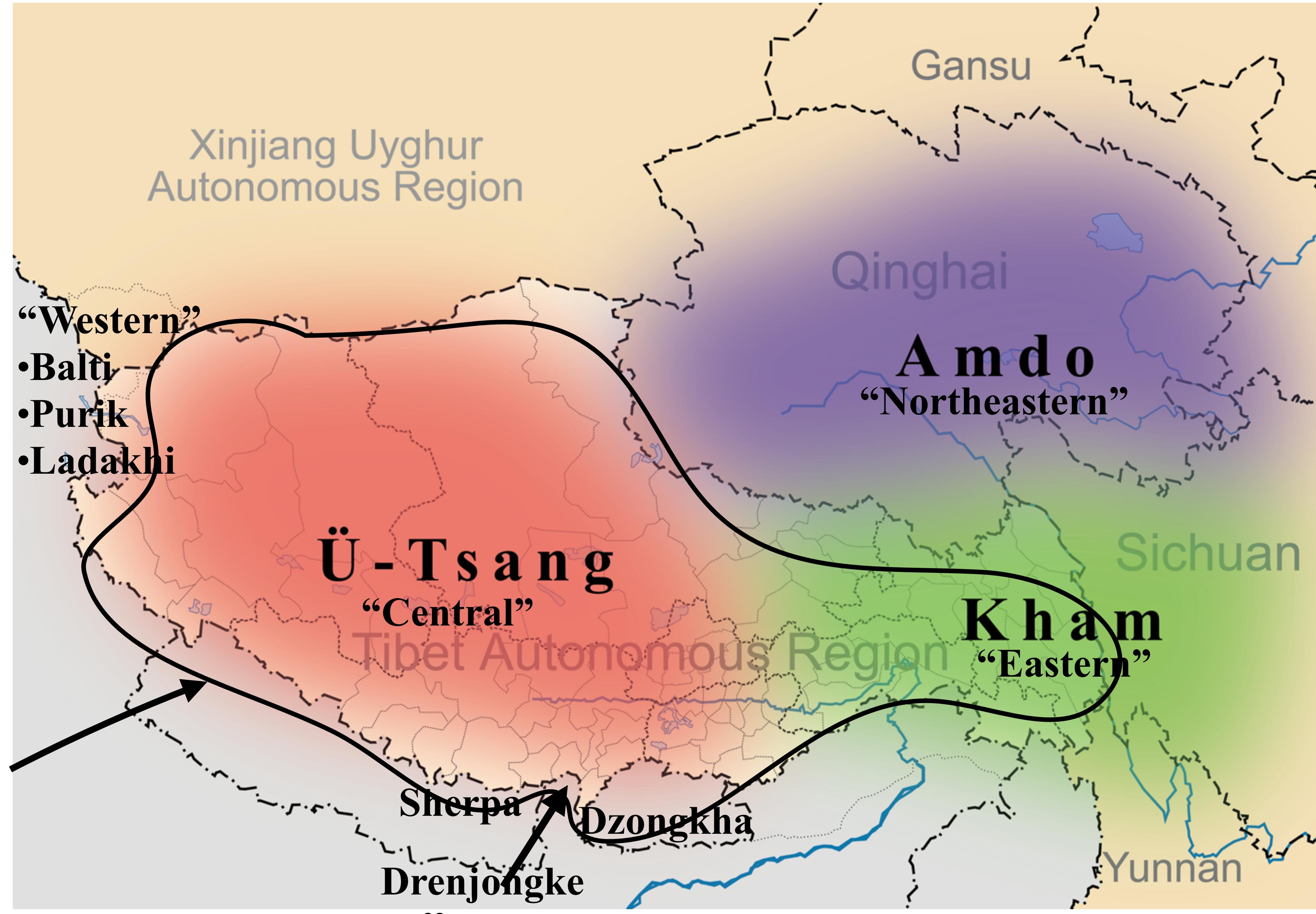
Tibetan dialects

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
<i>khrag</i>	[kʂʌk]	[t̪çɣy]	[t ^h ák] ([t ^h ák])	‘blood’
<i>rtswa</i>	[xst̪soa]	[xt̪sa]	[tsá]	‘grass’
<i>spyang ki</i>	[spjan̪.ku]	[xt̪can̪.kʰɣ]	[tʃán̪.gú]	‘wolf’
<i>bcu bdun</i>	[t̪cub.đun]	[t̪çɣb.đɣn]	[tʃúp.t᷑] ([tʃúp.t᷑])	‘seventeen’

(Adapted from Caplow 2013)

Tonogenesis

(tonal dialects only)

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

Laryngeal contrasts

	Etymological onsets							Innovative features
Orthography	ས	པ	ਬ	ڦ	س	ڙ	ڦ	
Old Tibetan	s ^ə pa	p ^h a	ba	s ^ə ba	sa	za	b ^ə za	aspiration allphonic
Northeastern and Western dialects	spa	p ^h a	ba ~ wa	বba	sa	za	za	cluster simplification aspirated/unaspirated contrast
Eastern dialects	pá	p ^h á	pà	bà	sá	zà	zà	tonogenesis cluster simplification
Central dialects (Lhasa)	pá	p ^h á	p ^h à	pà	sá	sà	sà	voiced clusters > voiceless voiced simplex > aspirated

Cross-linguistic evidence (after)

No tone,
no C-V lag

Arabic
Catalan
English
German
Georgian
Italian
Romanian

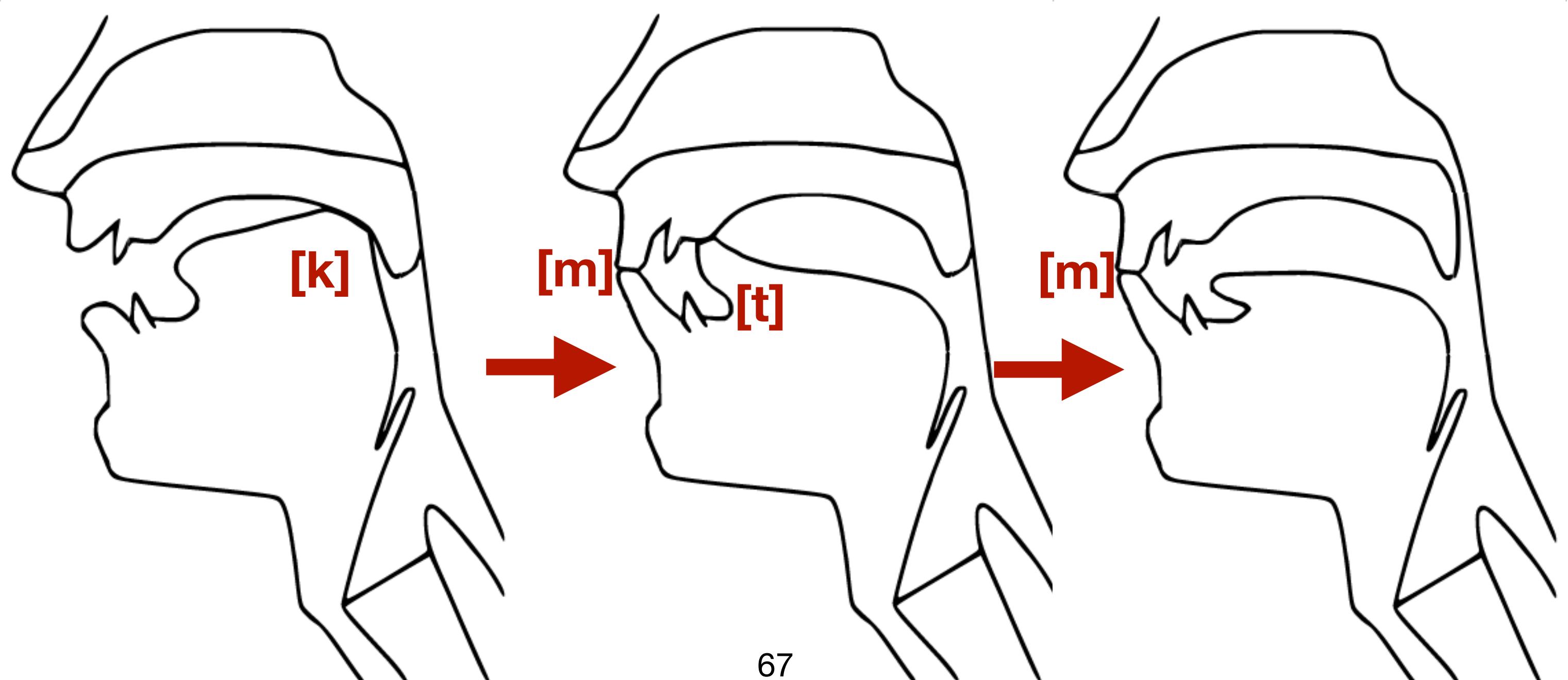
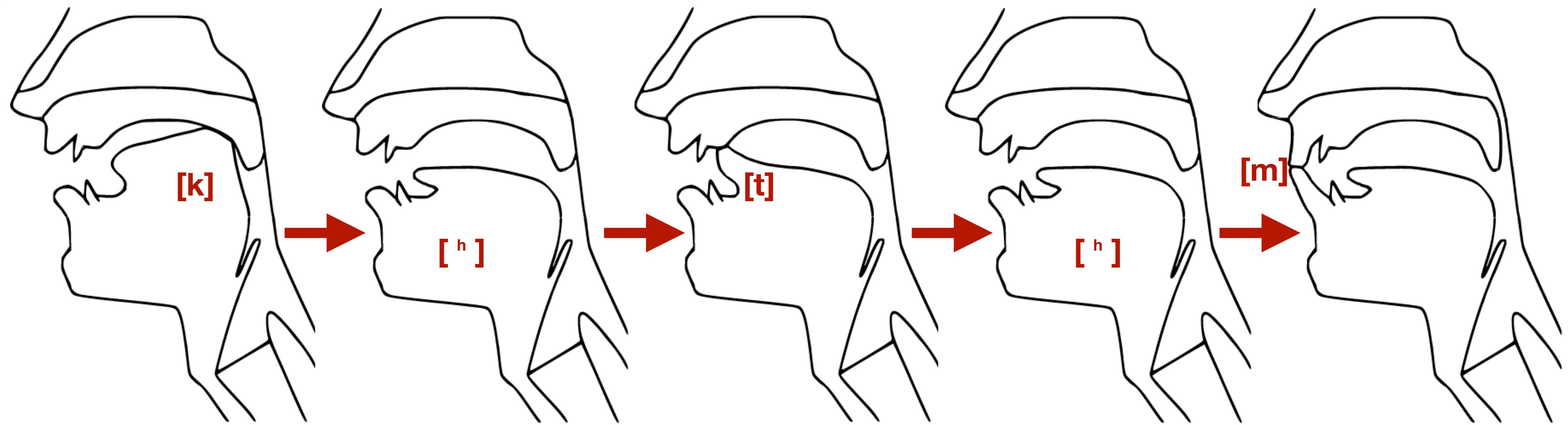
Tone

Swedish
Serbian

C-V lag

Mandarin
Thai
Tibetan

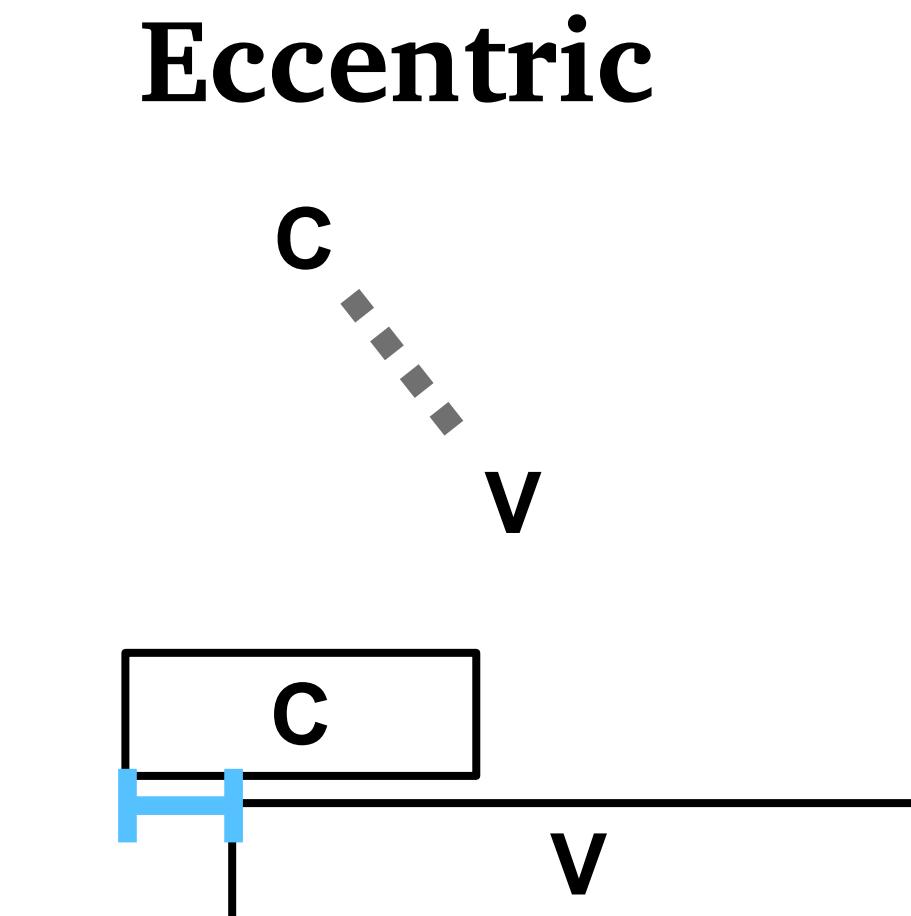
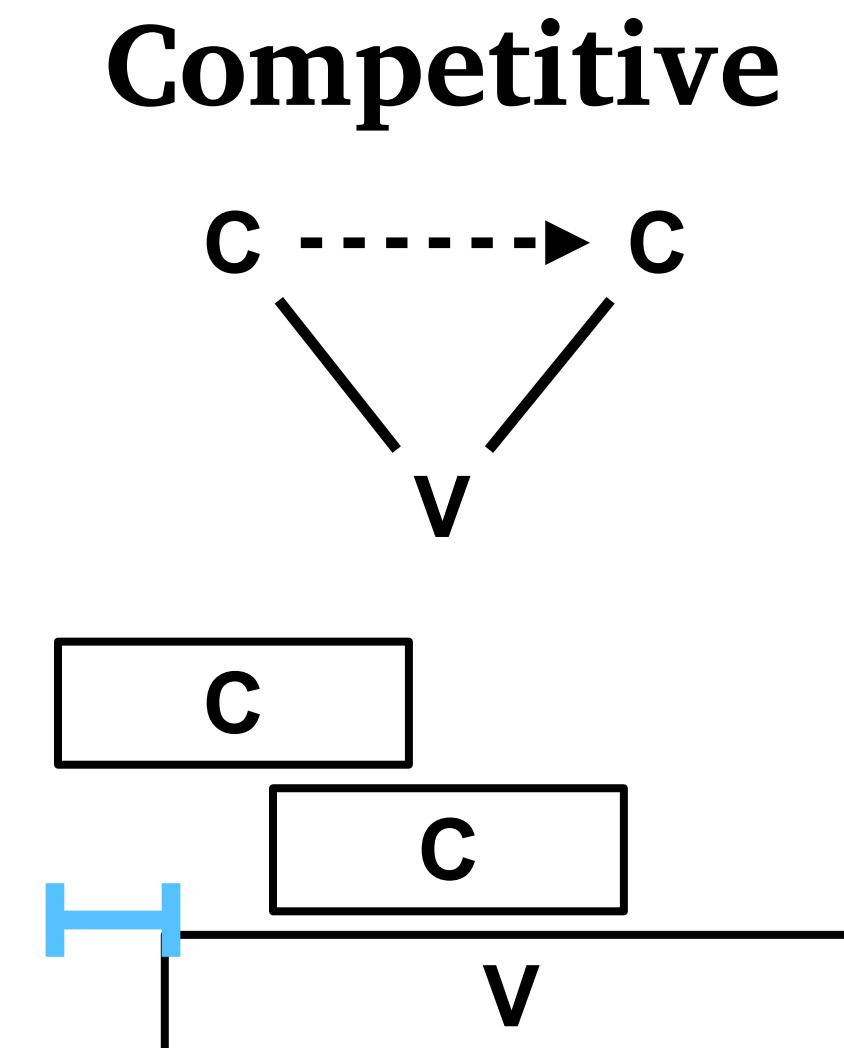
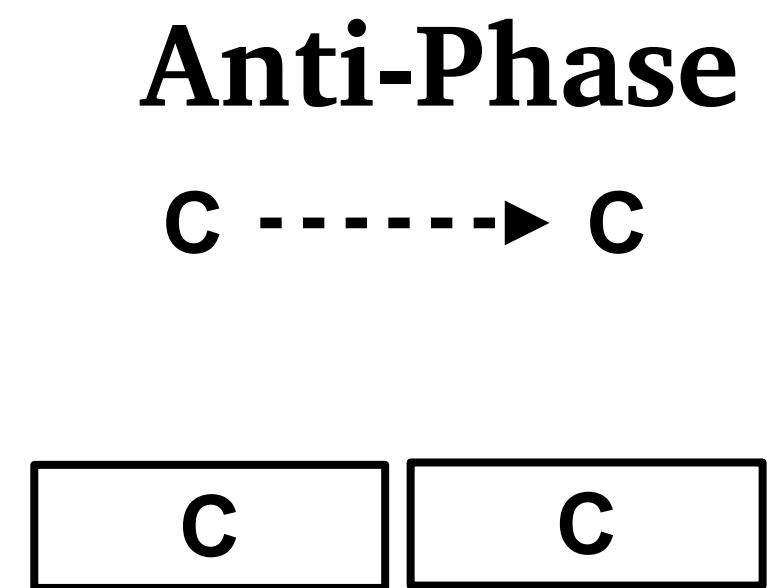
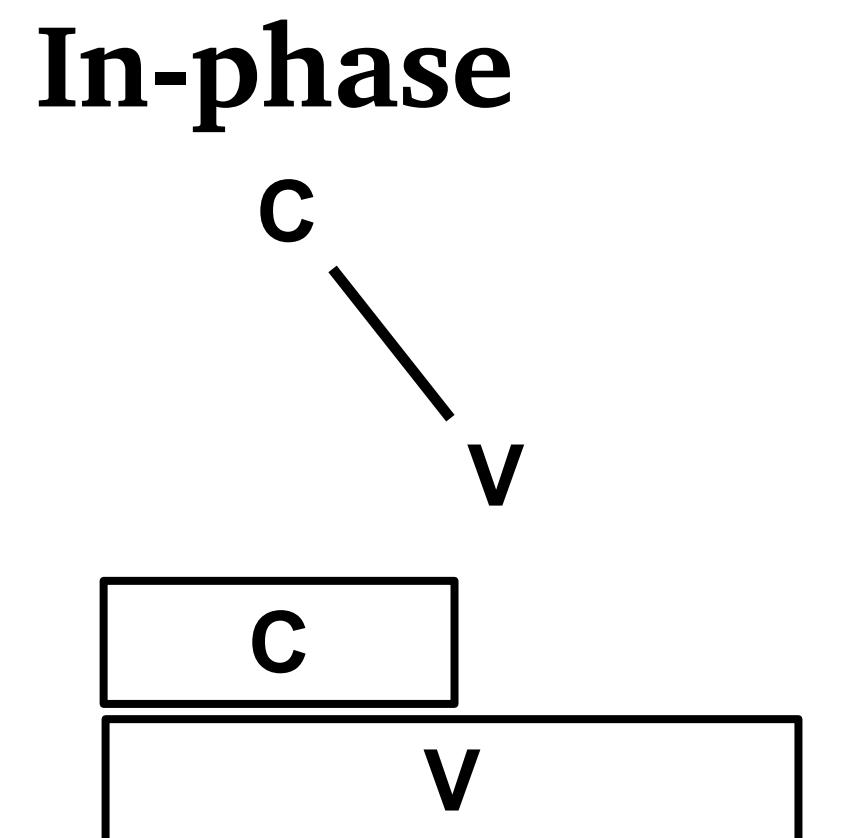
also Tibetan



[back to slide 7](#)

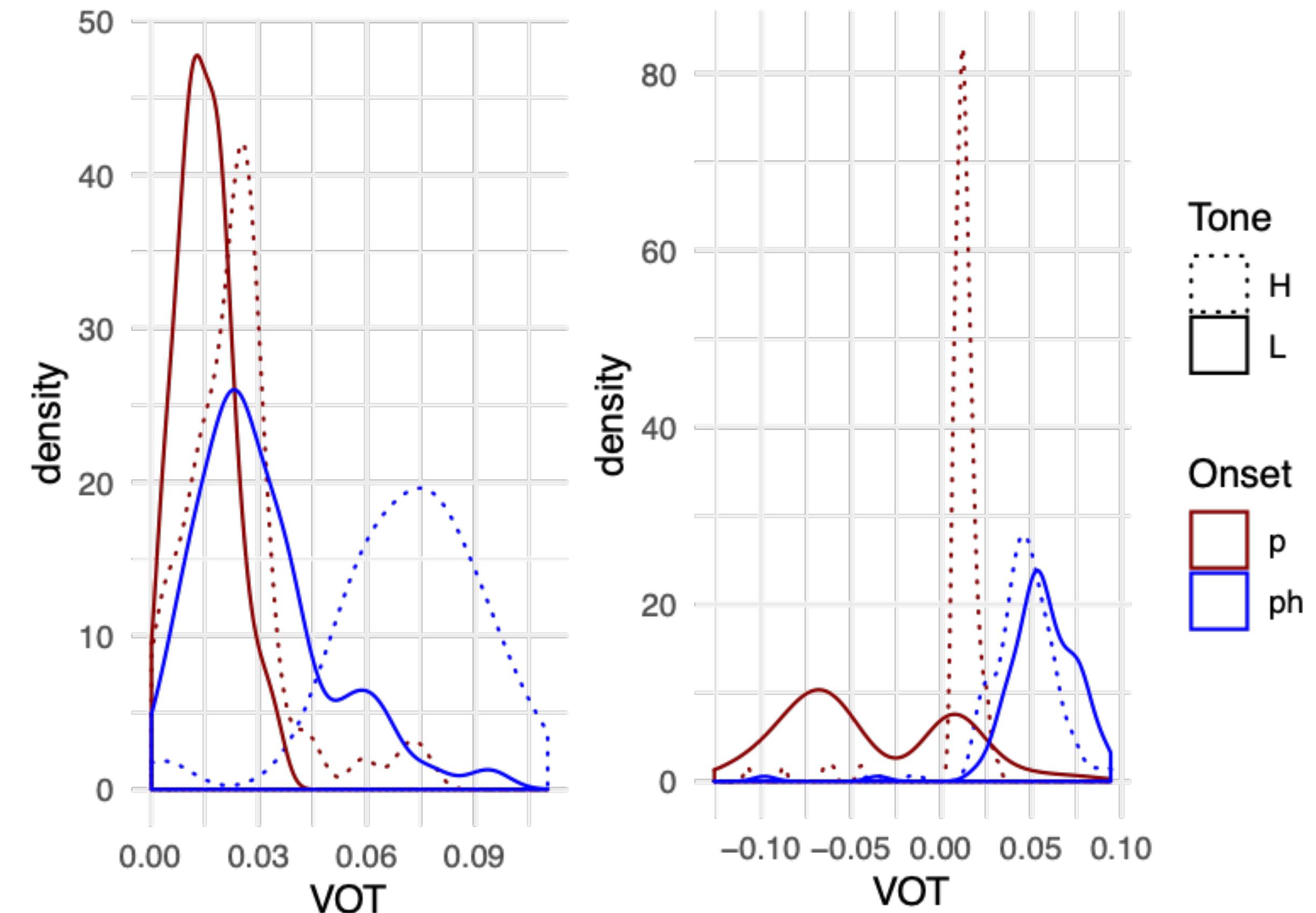
Coordinating gestures in time

- 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



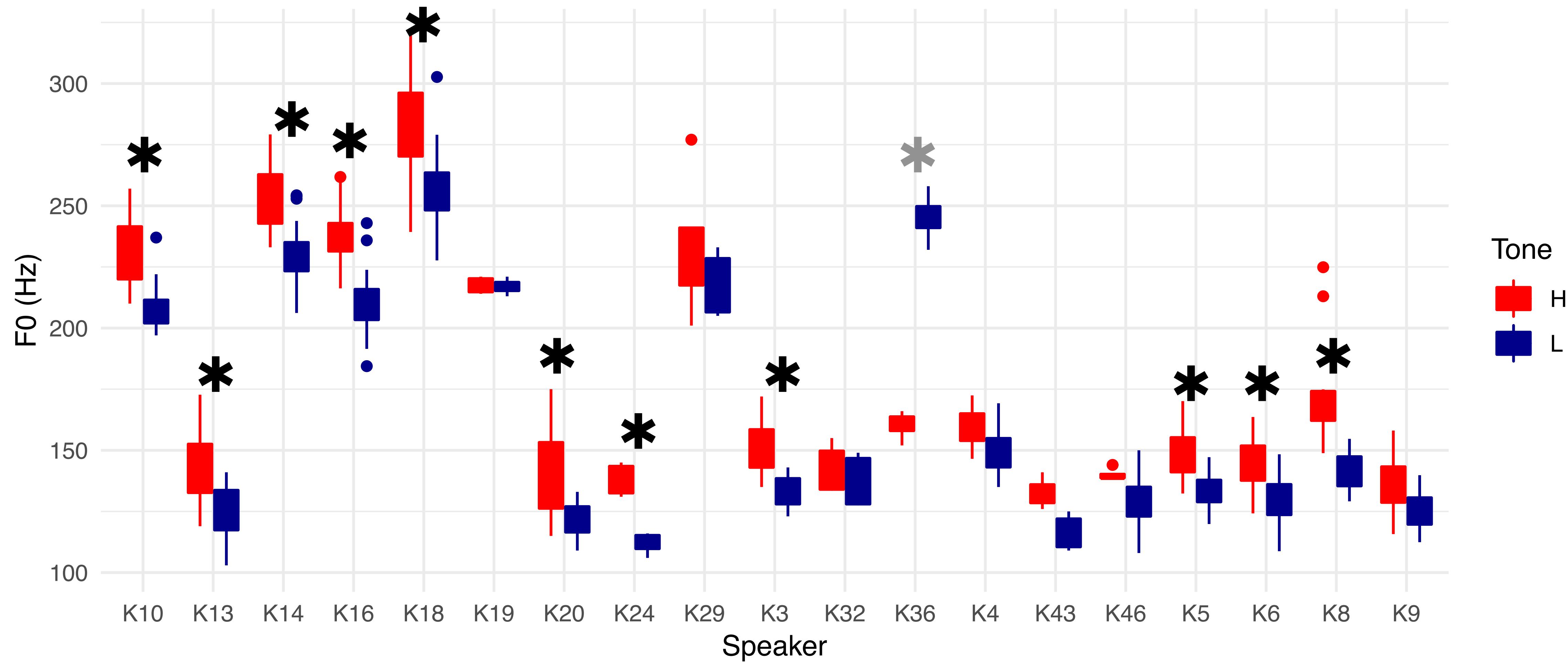
Two systems of laryngeal contrasts even in speakers with no F0 contrast (!!)

- 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

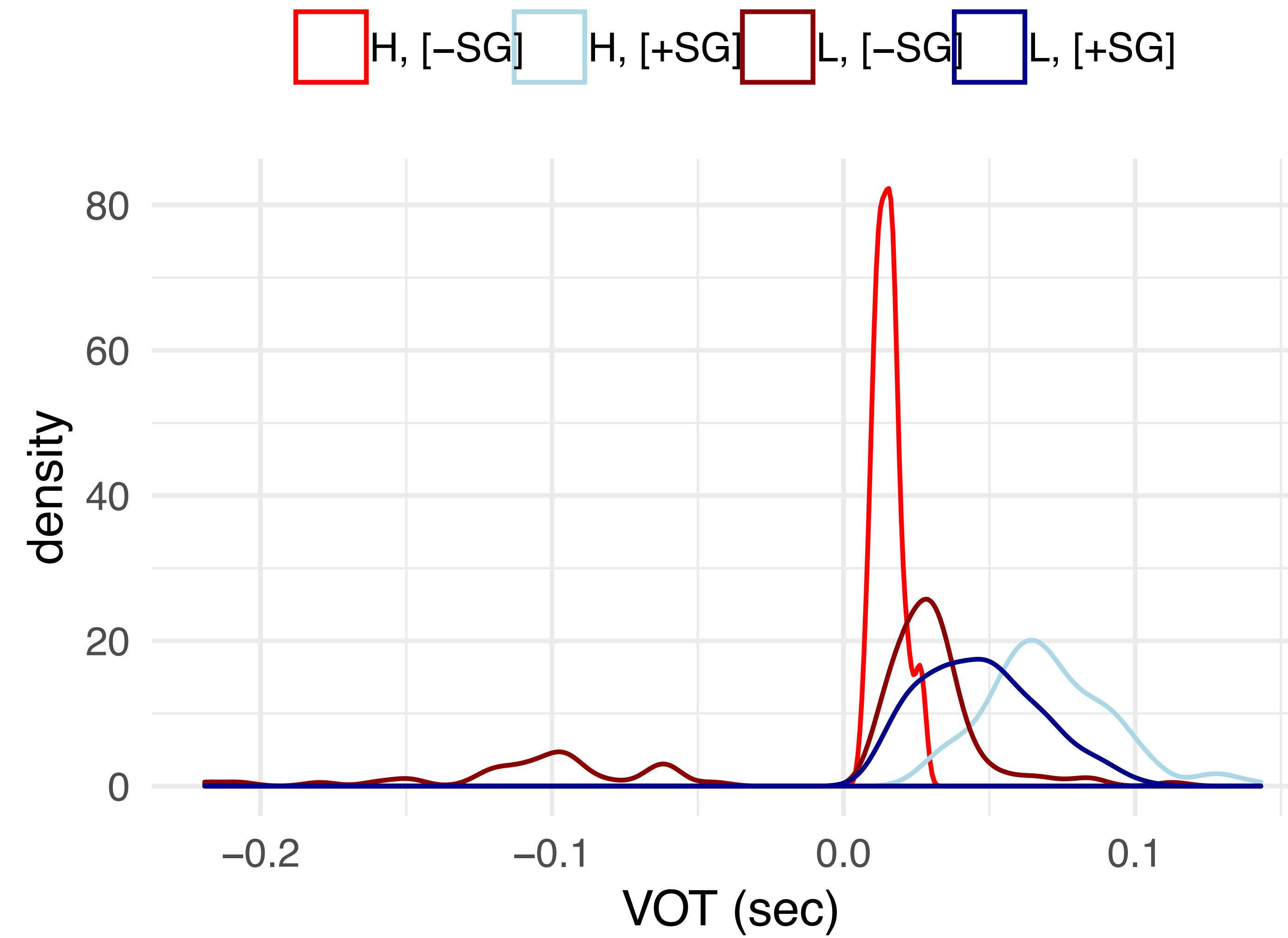


Does H have higher pitch than L?

Yes for 11/19, no for 7/19



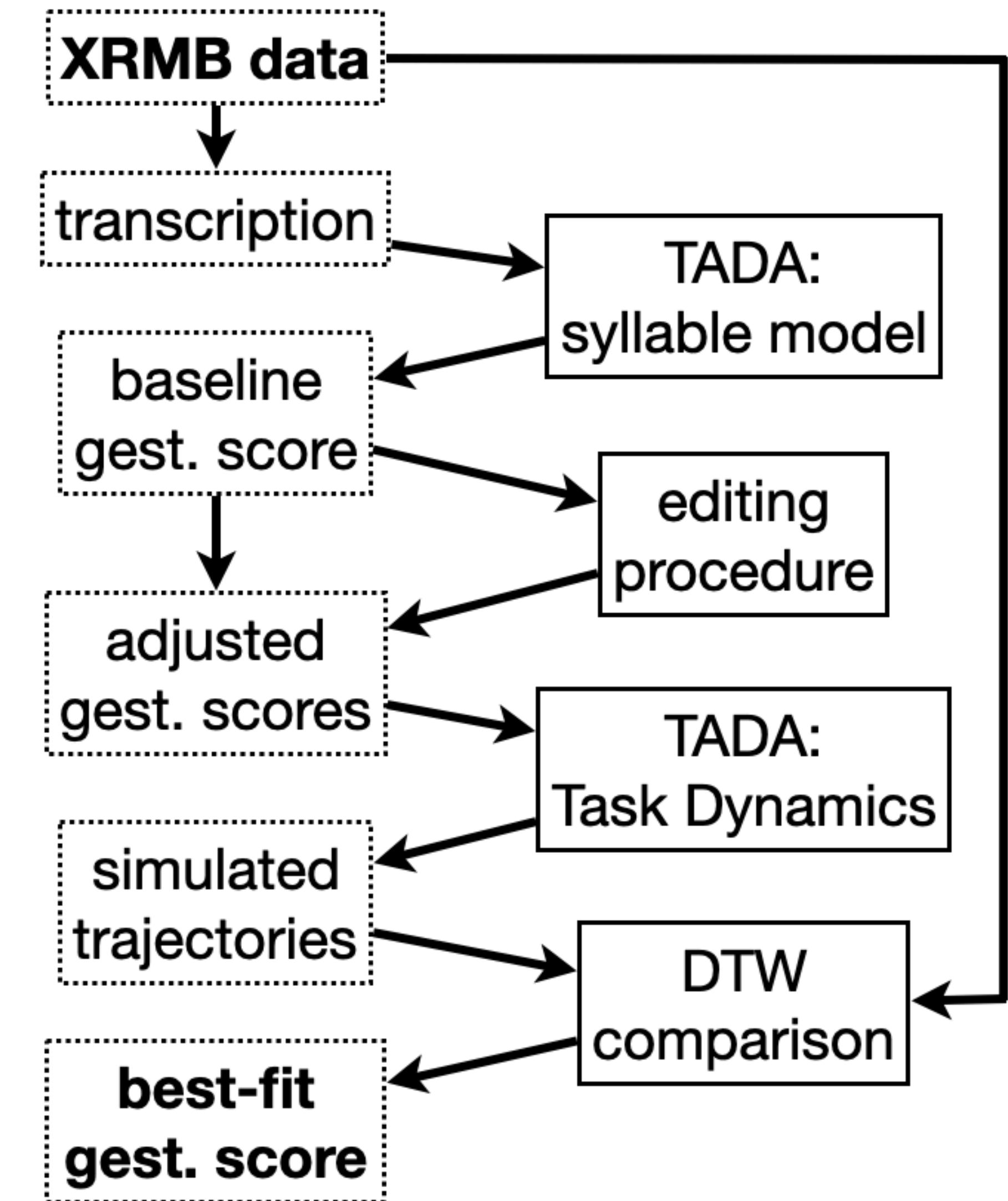
Consonant and tone categories



<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?