# Consistent C-V timing across speakers of diaspora Tibetan with and without lexical tone contrasts

Christopher Geissler[1], Jason Shaw[1], Mark Tiede[2], Fang Hu[3]

[1]Yale University, [2]Haskins Laboratories, [3]Chinese Academy of Social Sciences, Beijing

This material is based upon work supported by the National Science Foundation under Grant Number 1928750

# Introduction: Gestures

- Articulatory gestures: abstract, dynamic representations of controlled movements of the vocal tract

*(e.g. Browman & Goldstein 1986)*

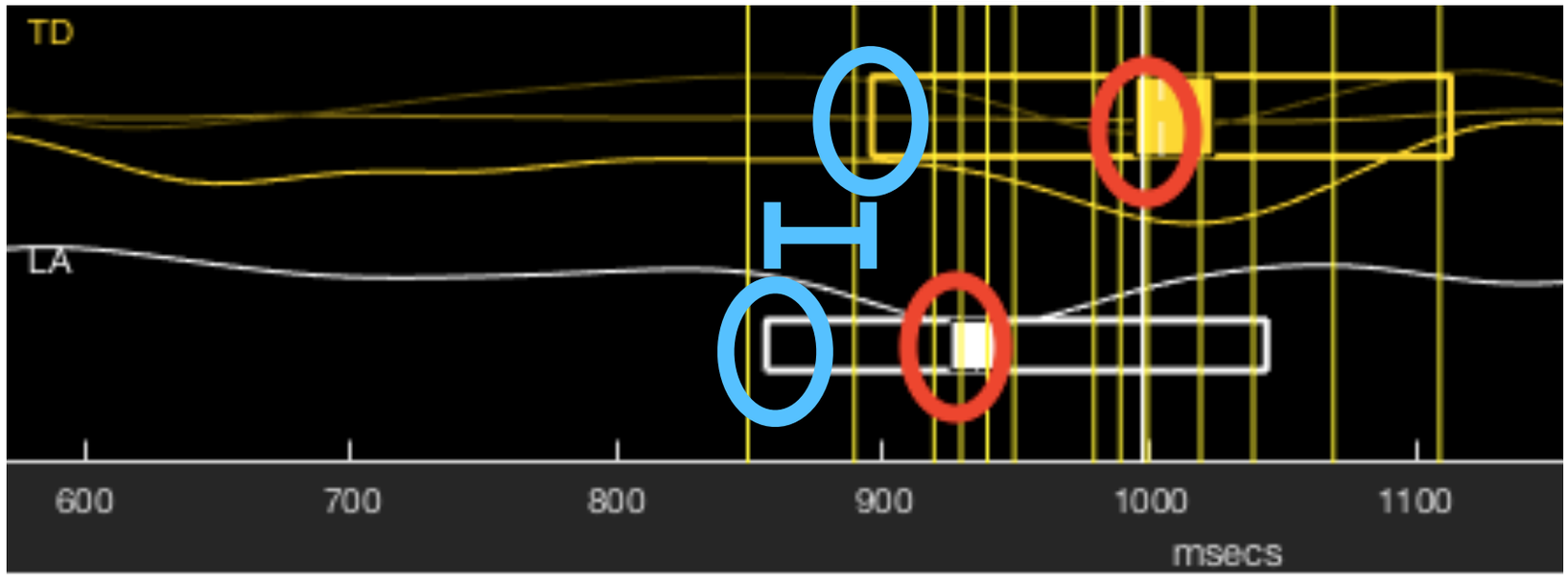
- How are gestures timed with each other?

[mu]

"C-V lag"

**Tongue Dorsum**: **front** ↓ **back**

**Lips**: **open** ↓ **closed**



# Questions

- How are speech gestures timed with each other?
- Which gestures are encoded in the phonological representation?
- How are those gestures coordinated?

# Which gestures?

- Strong Articulatory Phonology hypothesis: only gestures encoding phonological contrasts are represented

- C, V: specified for constriction location and degree

*(e.g. Browman & Goldstein 1986)*

- tones: relative F0 excursions

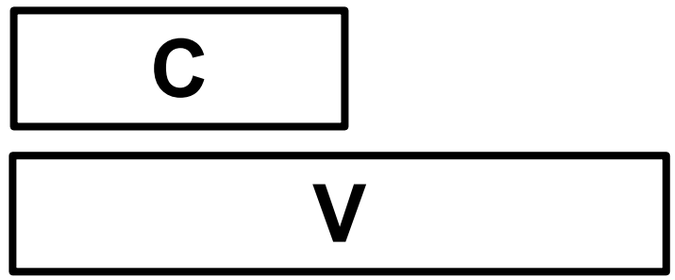
*(Gao 2008, Katsika et al 2014)*

# How are gestures coordinated?

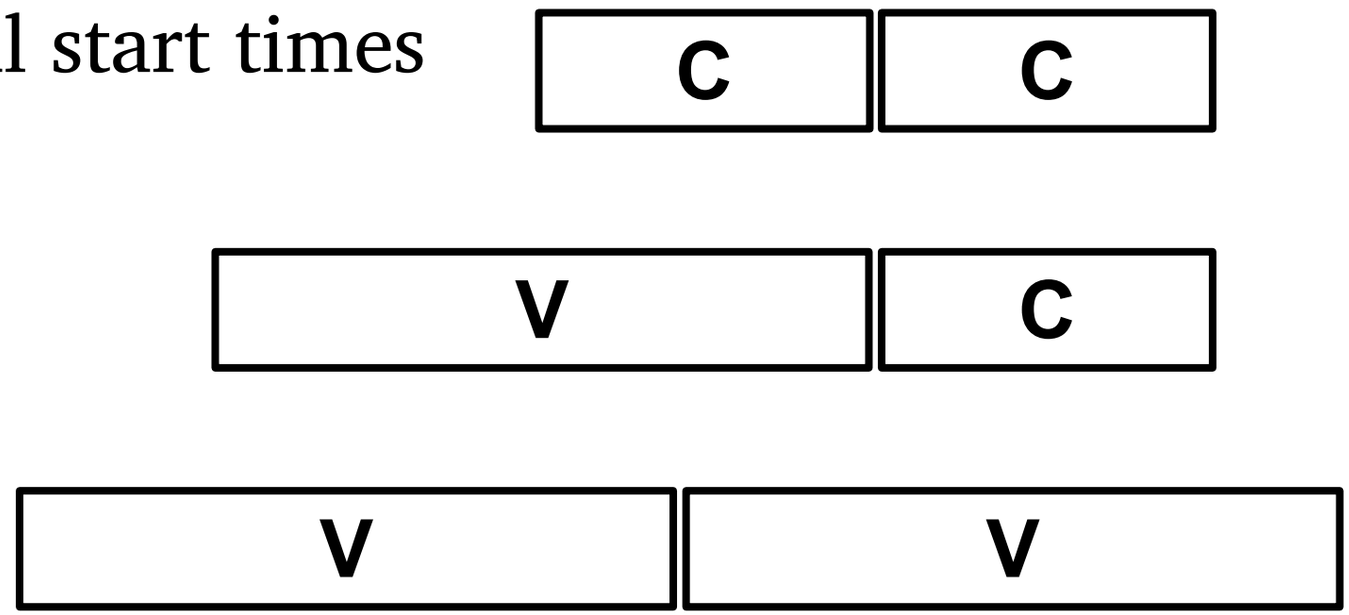
- *Planning oscillators* coordinate gestures with cyclic phasing
- In-phase (0°) and anti-phase (180°) coupling modes learned more easily, as in general motor coordination
  *(Browman & Goldstein 2000, Saltzman et al 2008)*
- Other phasing modes are possible, but more difficult (eccentric timing)
  *(Goldstein 2011)*

# Two stable coupling modes

- in-phase → synchronous start times

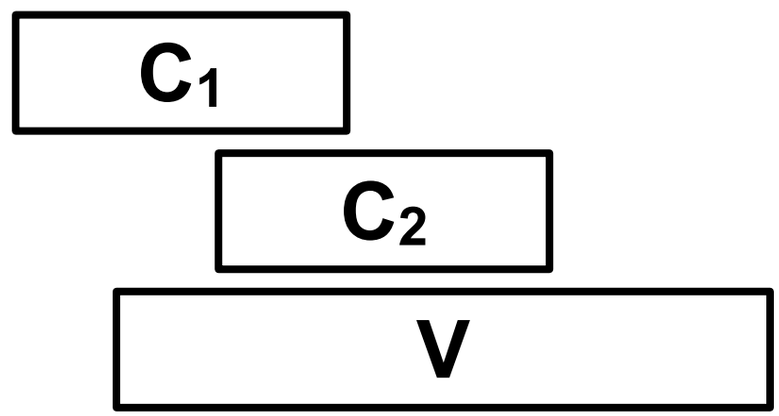



- anti-phase → sequential start times



# Exceptional C-V timing

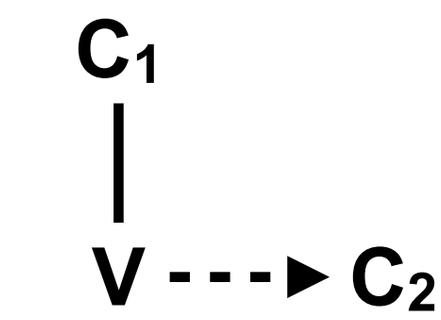
- Onset clusters often show partial overlap ("C-center")
  *(Browman & Goldstein 1988, inter alia)*
- Exceptional clusers:

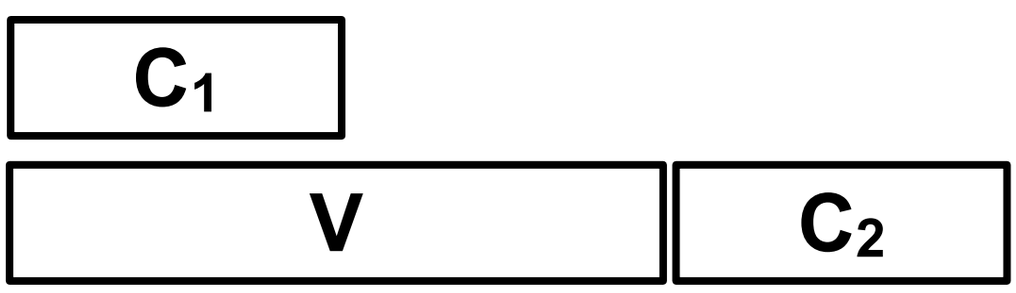


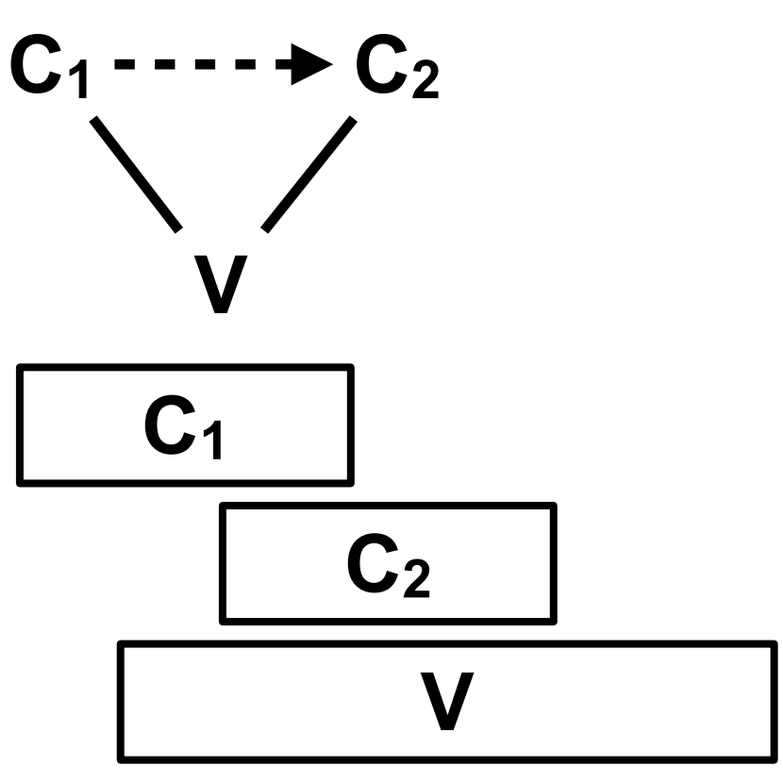

- some CV timing unchanged when add earlier C:
  - Italian /sC/ onsets *(Hermes et al. 2008, 2011)*
  - Moroccan Arabic *(Shaw et al 2009)*, Tashlhyit Berber *(Goldstein et al. 2007, Hermes et al. 2017)*

# Explanation: competitive coupling

- Account for partial overlap through competition beween in-phase and anti-phase coupling:
  *(Browman & Goldstein 2000, Nam & Saltzman 2003 )*



# C-V timing with tone

- Intonational tone doesn't affect C-V timing in some languages: German and Italian *(Niemann et al. 2011)*, Catalan *(Mücke et al. 2012)*. C-V lag $<$10ms

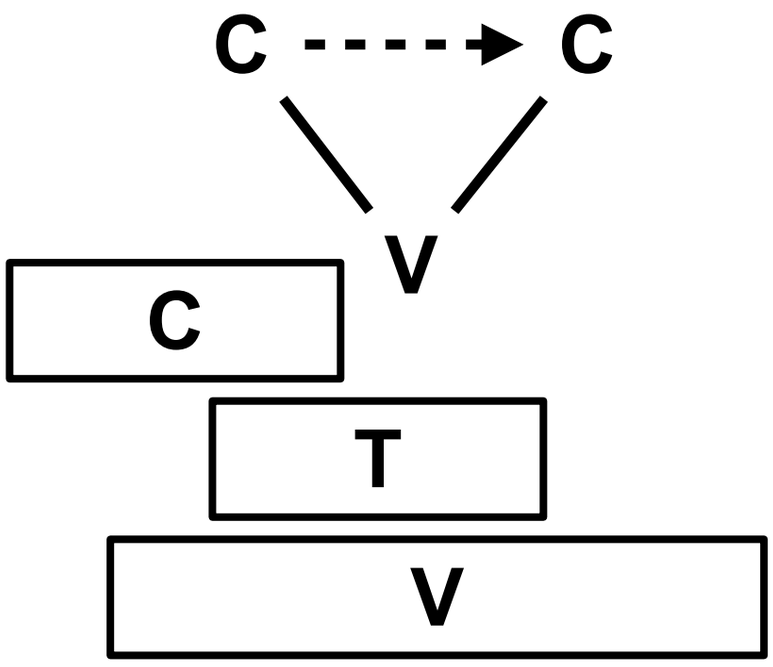
- Longer C-V lag in Mandarin *(Gao 2008)*, Thai *(Karlin 2014)*, and Lhasa Tibetan *(Hu 2016)*. C-V lag$\sim$50ms

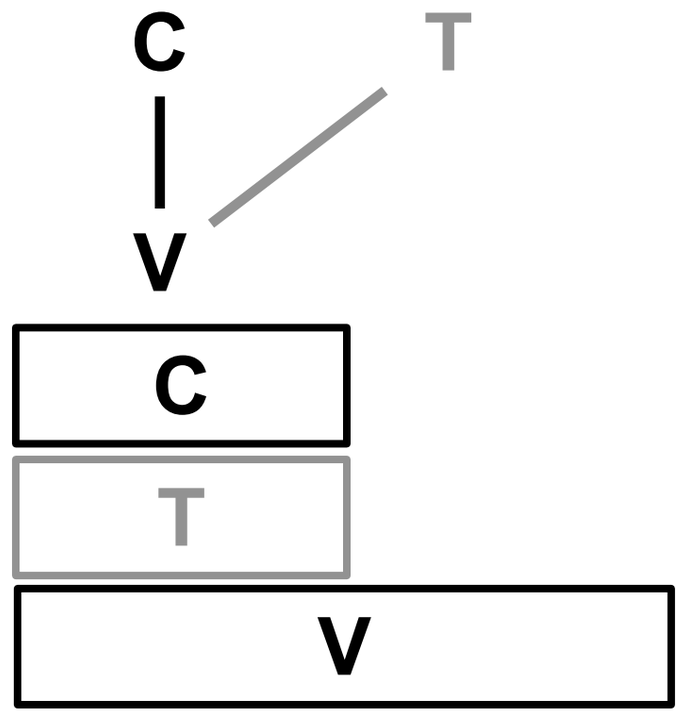
- Toneless syllables in Mandarin show reduced C-V lag relative to their fully-tonal counterparts *(Zhang et al. 2019)*

# C-V lag and tone

- Different potential structures for tone gesures:

| Mandarin | Italian |
| --- | --- |
| Thai | German |
| Tibetan | Catalan |



# Evidence for tone ~ C-V lag

- Tone is associated with longer C-V lag:

  - in lexical tone languages

    vs. non-lexical tone languages

  - in tonal and toneless syllables

    in the same language

  - present study: across speakers with

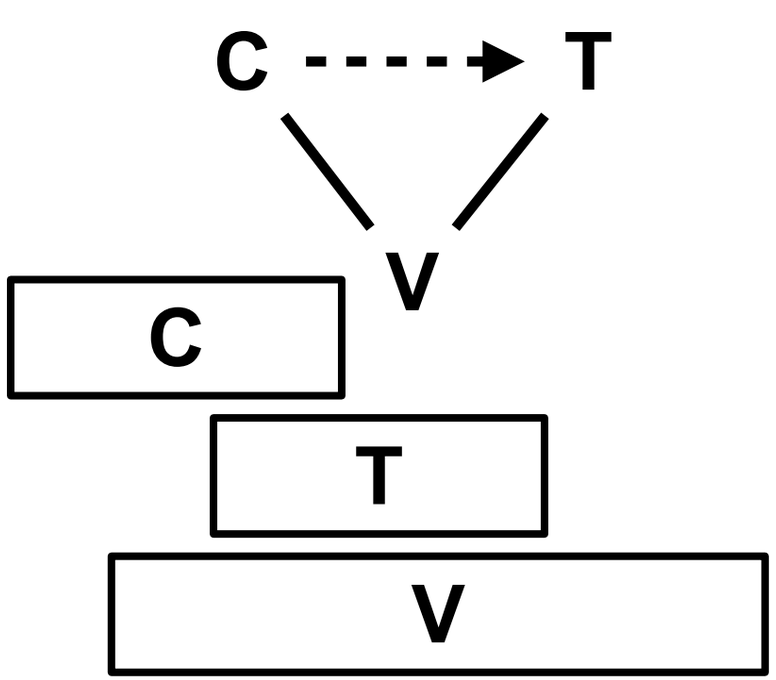
    vs. without tone contrast in the same language

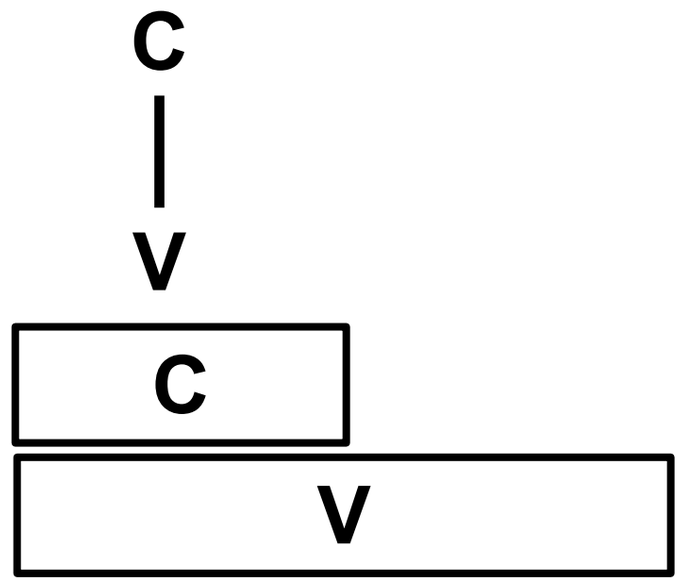
# Hypothesis

- In a language where some speakers produce a tone contrast and others do not:

tone-contrasting speakers → positive C-V lag

non-tone-contrasting speakers→ near-zero C-V lag



# Tibetan

- Tonal and non-tonal dialects
  - Tone contrast: H vs. LH
    *(Duanmu 1992, Tournadre and Dorje 2003)*
- Speakers raised in post-1959 diaspora (India, Nepal) exposed to mixed input, acquire mixed features
  *(Geissler 2018)*

# EMA Experiment

- Electromagnetic Articulography (EMA) to track fleshpoints on the lips and tongue, alongside audio
  - goal: quantify timing of oral gestures
- 6 Tibetan speakers (4 female) raised in Diaspora
  - all multilingual, extensive dialect contact

# Methods

- Speakers read words in carrier phrase on a screen, in Tibetan orthography

- EMA sensors on each lip and three on tongue; head movement corrected w/r/t/ three sensors on rigid points of the head

- Gesture start labelled at 20% of peak velocity to target

# Stimuli

- Bilabial onsets: separate C and V articulators
- Back vowels following front vowel in consistent frame sentence
- /m p $p^h$/ * 2 tones * /a o u/ * CV/CVC syllables * mono/disyllabic * 10 repetitions

# Identifying tone contrast

- measured F0 at ten time-normalized points along [mV] syllables (60 words per speaker)

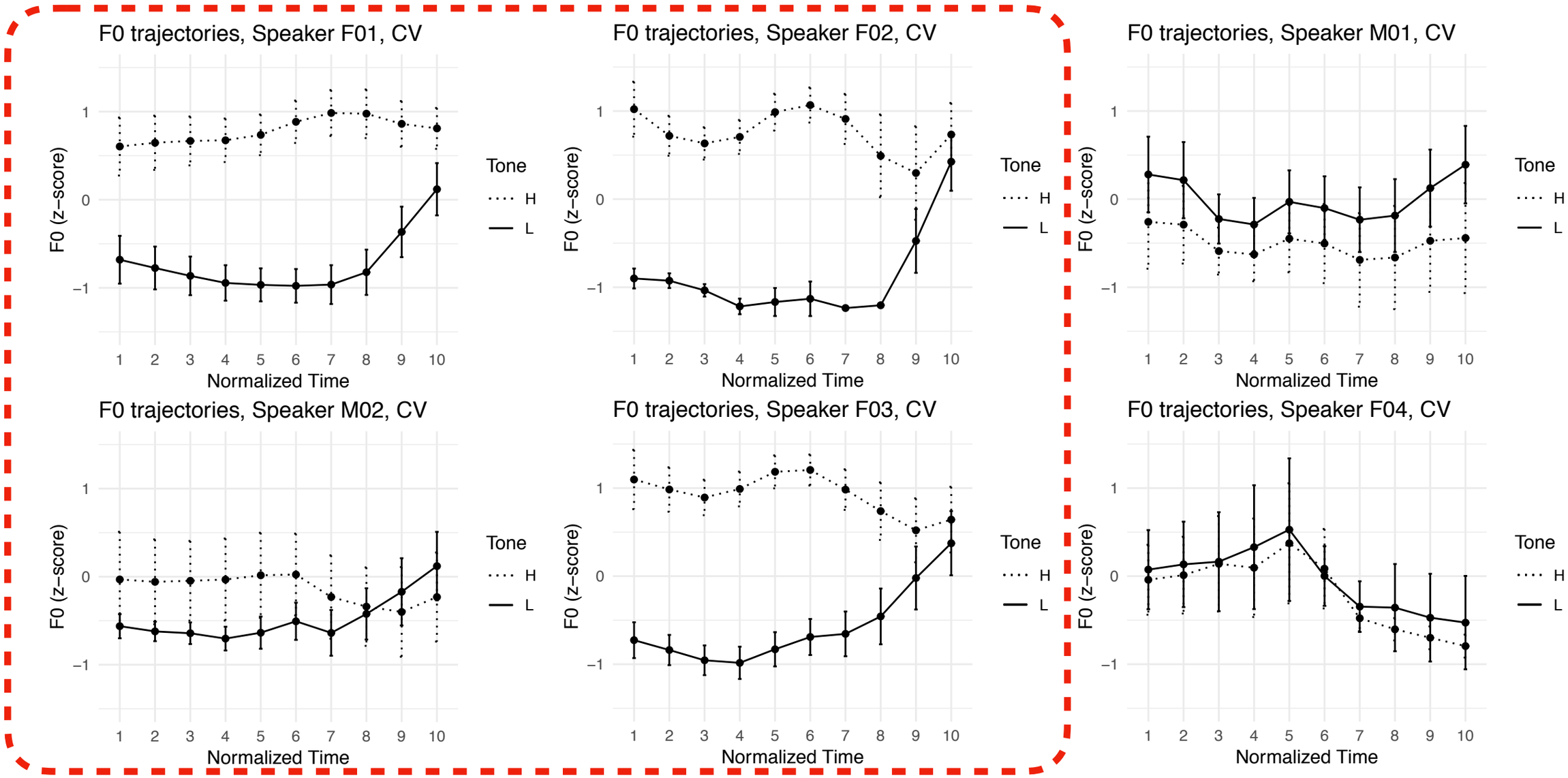
- acoustic analysis in Praat *(Boersma and Weenink 2018)*; VOT and time-normalized pitch calculated using Praat scripts *(DiCanio 2011, 2018)*

# Results: tone contrast

- fitted GAMMs to predict F0 based on:
    - parametric term for tone
    - smooth for timestep at reference value for tone
    - difference smooth across tones
    - random smooths by word

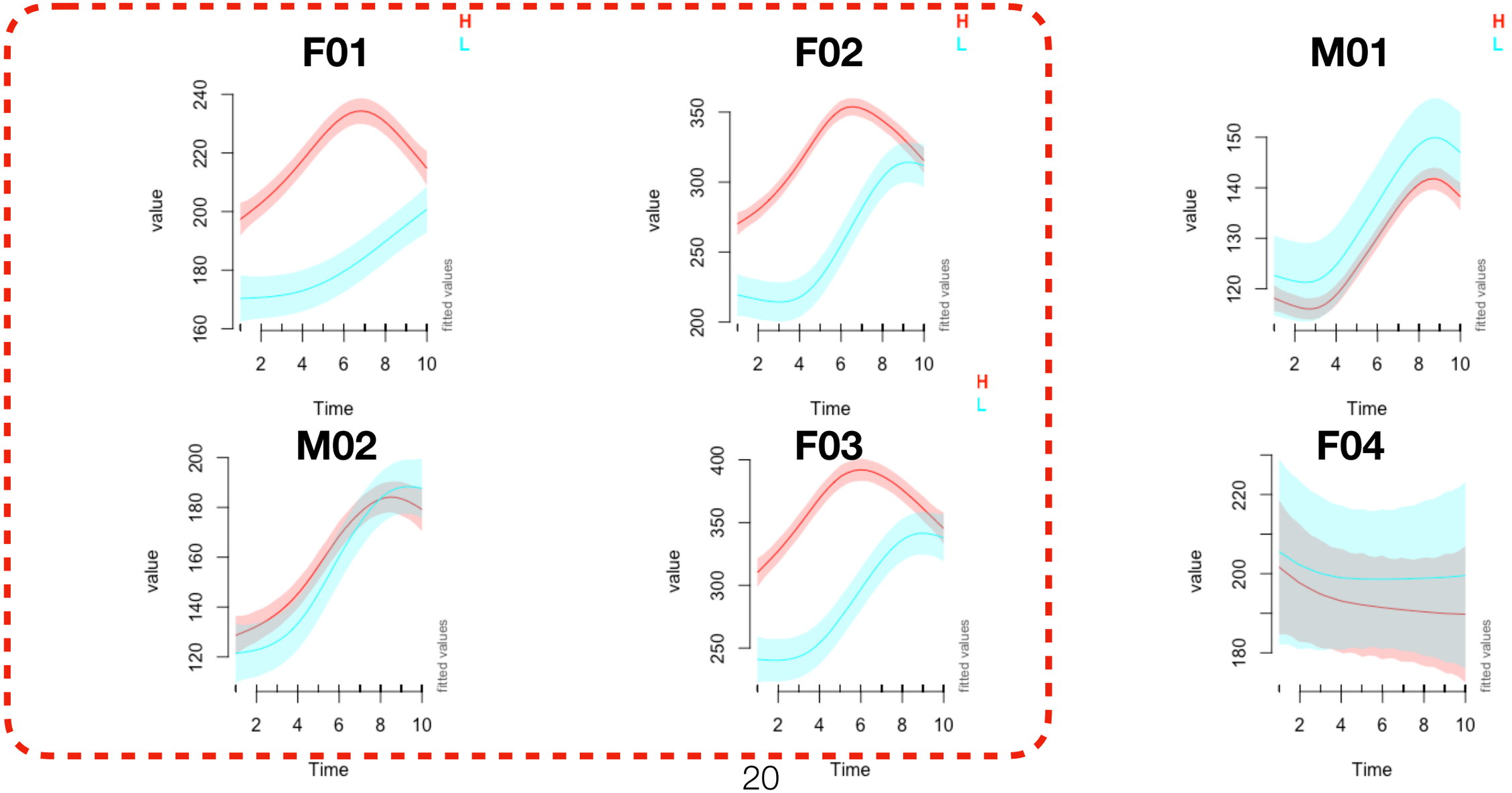
# Results: tone contrast

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



# Results: tone contrast

- Confirm with GAMM (smooths for tone plotted)

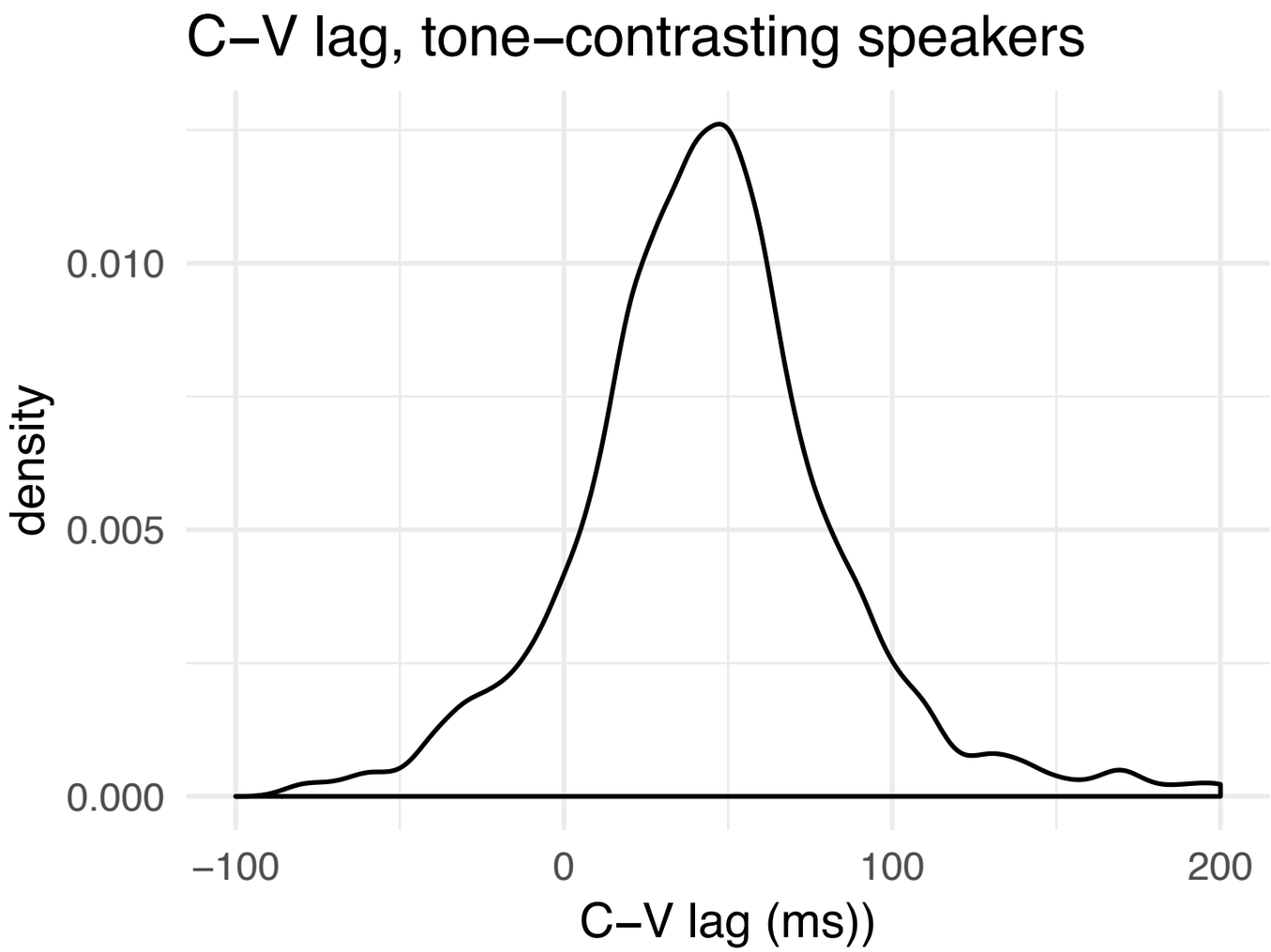


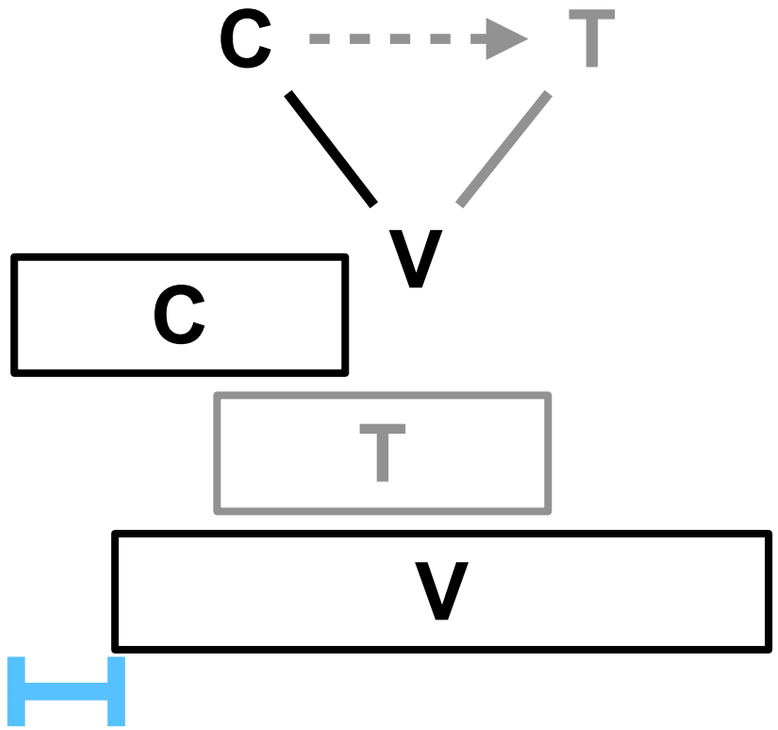
# Results: tone contrast

| term | F01 | F02 | F03 | M02 | M01 | F04 |
|---|---|---|---|---|---|---|
| tone (parametric) | * | * | * | | | |
| time smooth | * | * | * | * | * | |
| difference smooth by tone | * | * | * | * | | |
| random smooths by word | * | * | * | * | * | * |
| Deviance explained | 67.8% | 94.8% | 80% | 71.6% | 77% | 8.57% |

# Results: C-V lag

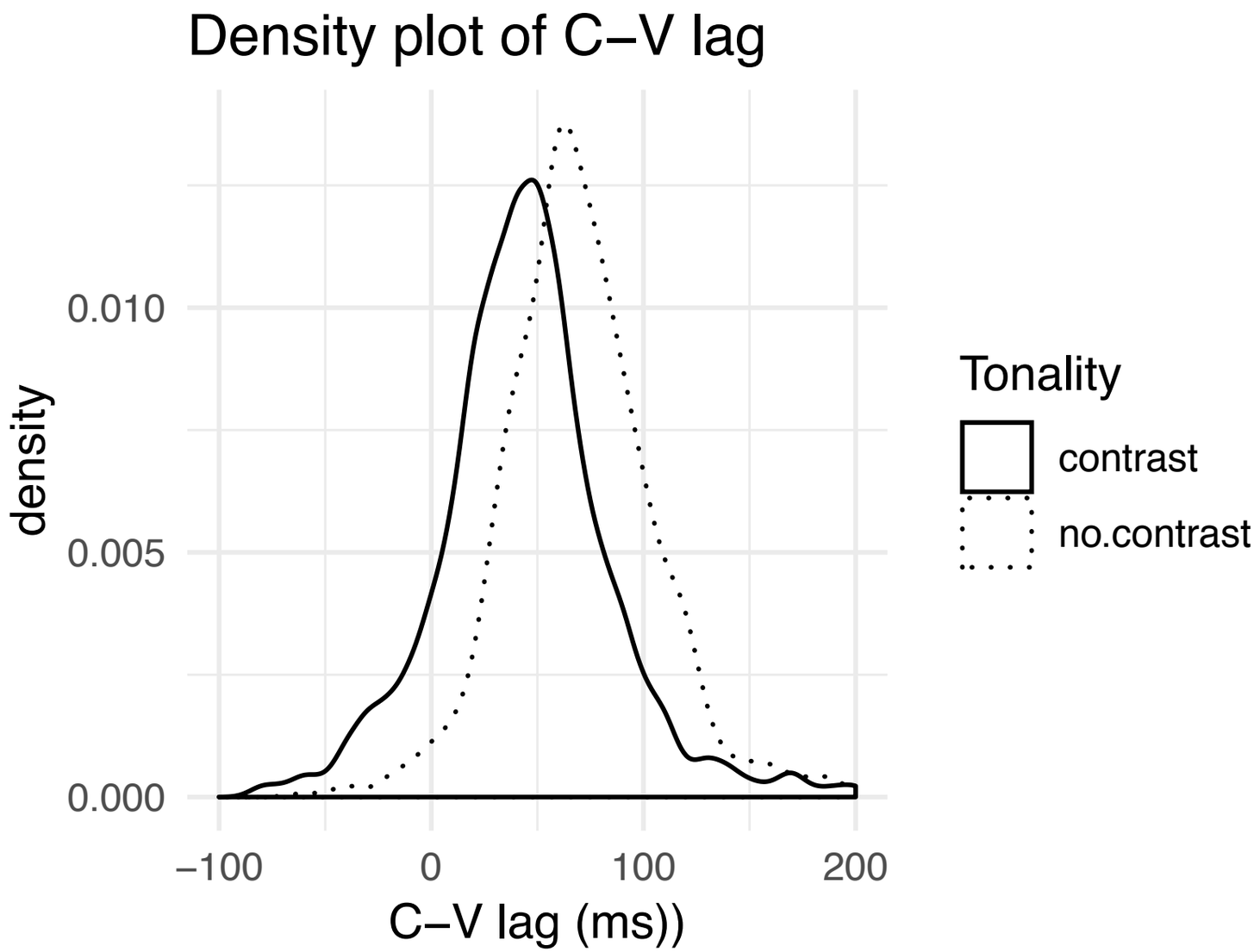
- Among tone-contrasting speakers, C-V lag is positive!



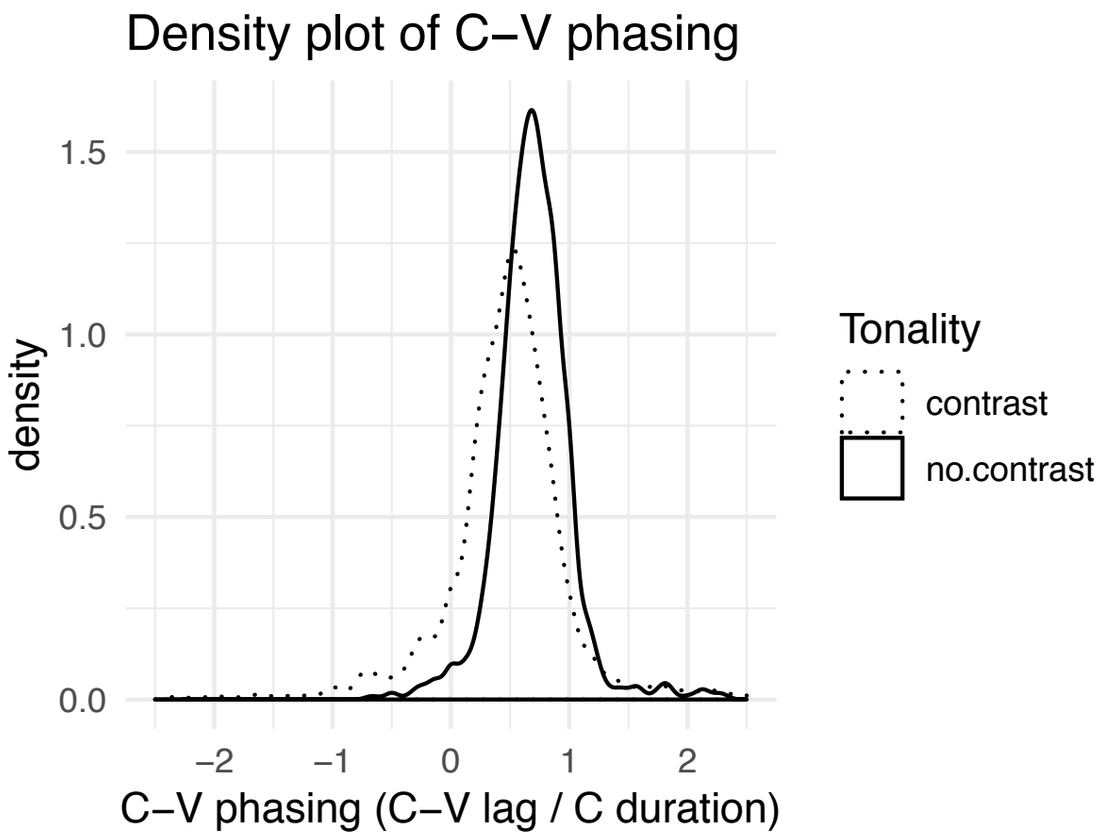
# Results: C-V lag

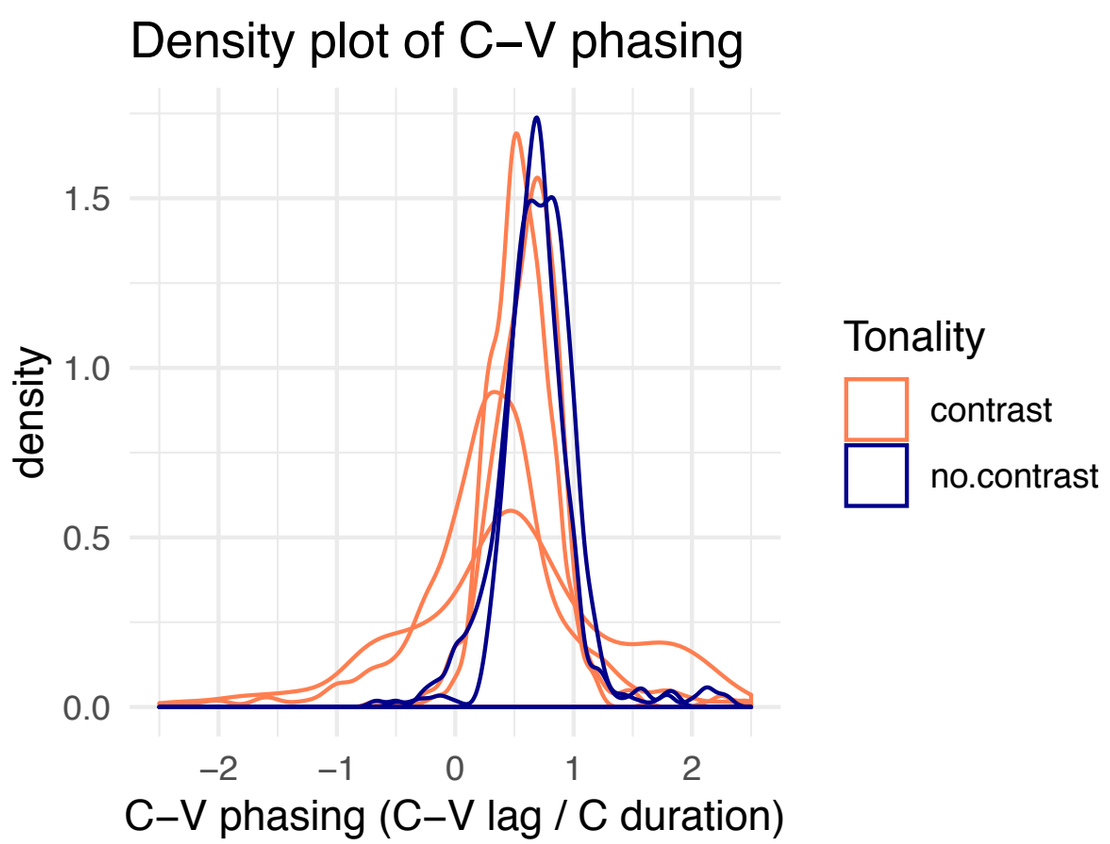
- … but also for non-contrasting speakers



# Results: C-V phasing

- C-V lag relative to C duration also similar for speakers with and without tone contrast

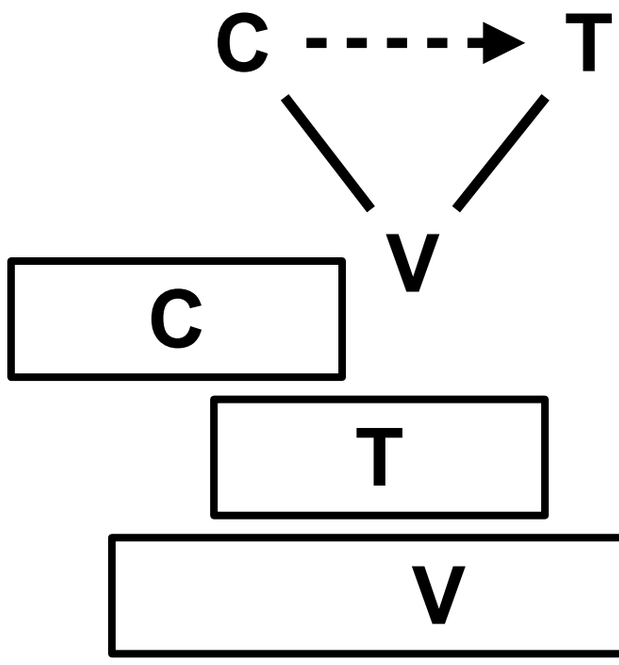
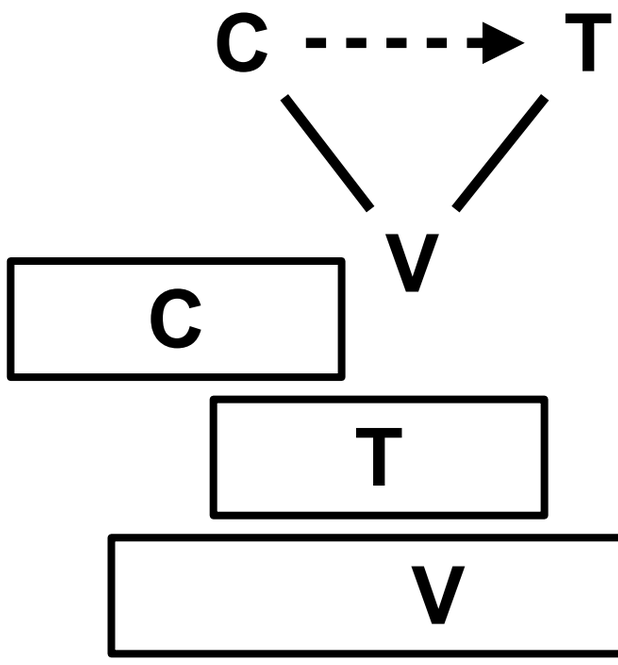
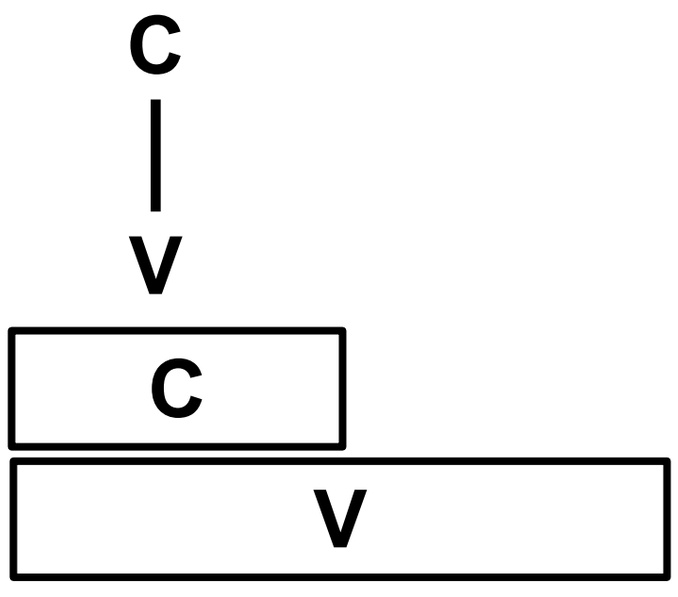
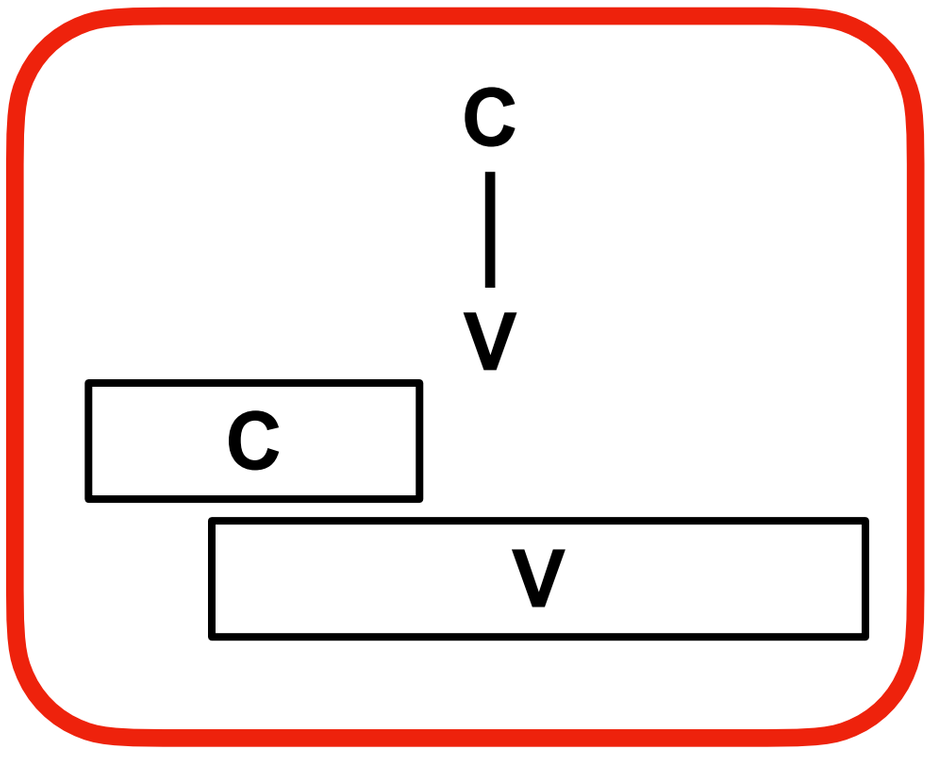


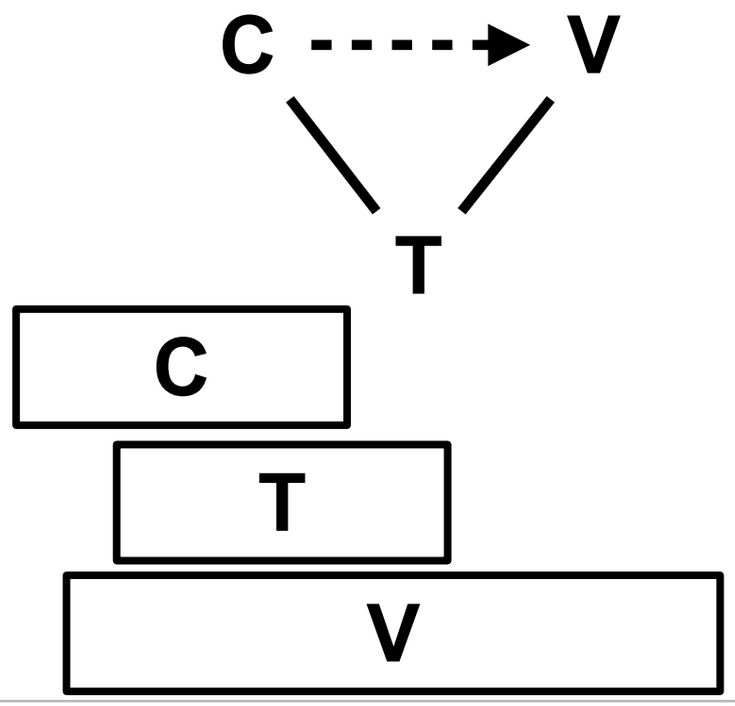
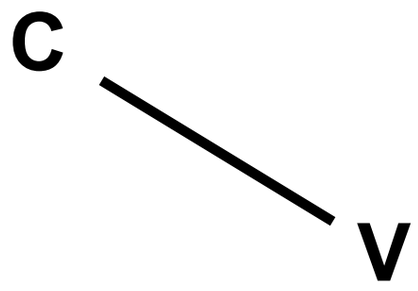
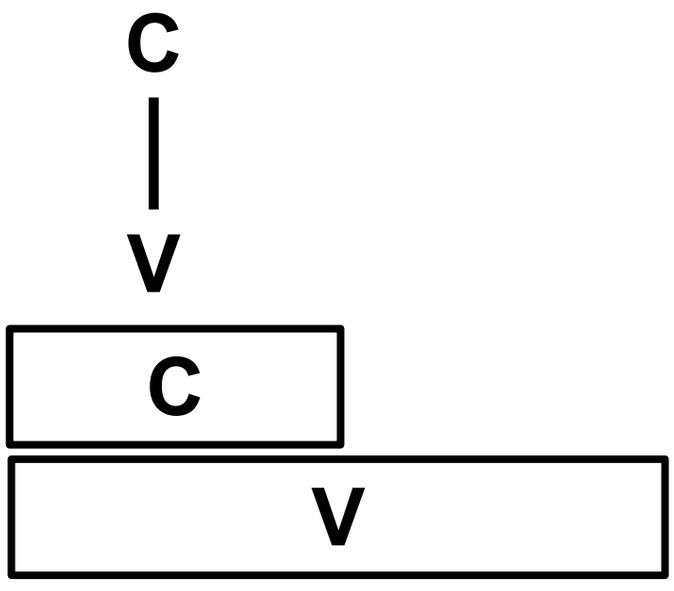
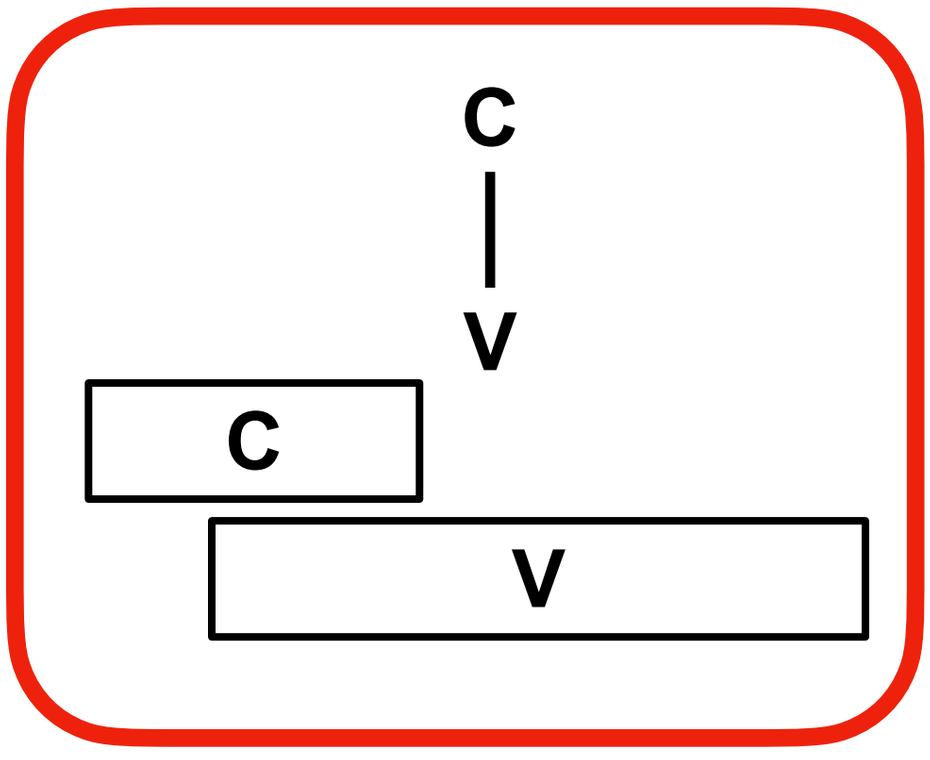



# Results: C-V phasing

- Confirmed results with comparison of LMMs:
    - baseline model: fixed effect of onset, random effects of speaker and word
    - comparison: baseline plus fixed effect of tone contrast

| model | Df | AIC | logLik |
|---|---|---|---|
| baseline | 6 | 12461 | -6224.5 |
| comparison | 7 | 12462 | -6223.7 |

| | Predicted | Observed |
|---|---|---|
| Tone contrast |  |  |
| No tone contrast |  |  |

| | Predicted | Observed |
|---|---|---|
| Tone contrast |  |  |
| No tone contrast |  |  |

# Summary

- Some diaspora speakers contrast tone, others don’t
- Observe long C-V lag in Tibetan, like Thai and Mandarin
- … but speakers with and without tone show similar C-V lag
    - C-center-like timing learned even by speakers lackng the tone contrast

# Interpretation

- How to account for similar C-V lag across speakers with and without tone production contrast?

- Possibility 1: Non-contrasting speakers have a non-contrastive (tone?) gesture

- Possibility 2: Non-contrasting speakers use eccentric C-V timing

# Interpretation

- How account for similar C-V lag across speakers with and without tone production contrast?
- Possibility 1: Non-contrasting speakers have a non-contrastive (tone?) gesture
    - different from “Strong AP hypothesis” where only contrastive gestures are in the coupling graph

# Interpretation

- How account for similar C-V lag across speakers with and without tone production contrast?

- Possibility 2: Non-contrasting speakers learn the same C-V timing spoken around them

  - without competitive coupling, as eccentric timing

*(e.g. Marin & Pouplier 2010, Goldstein 2011)*

# Conclusions

- Tibetan speakers with and without a tone production contrast showed similar C-V lag

- Speakers can learn eccentric timing relations resembling those of other members of the speech community

  - Eccentric timing can resemble competitive coupling

# References I


Browman, Catherine P. & Louis M. Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3. 219–252.

Browman, Catherine P. & Louis Goldstein. 1988. Some notes on syllable structure in articulatory phonology. *Phonetica* 45(2–4). 140–155.

Browman, Catherine P. & Louis Goldstein. 2000. Competing constraints on intergestural coordination and self-organization of phonological structures. *Les Cahiers de l'ICP. Bulletin de la communication parlée* (5). 25–34.

Duanmu, San. 1992. An Autosegmental Analysis of Tone in Four Tibetan Languages. *Linguistics of the Tibeto-Burman Area* 15:1.

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

Goldstein, Louis, Ioana Chitoran & Elisabeth Selkirk. 2007. Syllable structure as coupled oscillator modes: evidence from Georgian vs. Tashlhiyt Berber. *Proceedings of the XVIth international congress of phonetic sciences*, 241–244.

Goldstein, Louis. 2011. Back to the Past Tense in English. In Rodrigo Gutiérrez-Bravo, Line Mikkelsen & Eric Potsdam (eds.), 69–88. Santa Cruz, CA: Linguistics Research Center, UC-Santa Cruz Department of Linguistics.

Hermes, Anne, Martine Grice, Doris Mücke & Henrik Niemann. 2008. Articulatory indicators of syllable affiliation in word initial consonant clusters in Italian. *Proceedings of the 8th International Seminar on Speech Production* 433–436.

Hermes, Anne, Doris Mücke & Bastian Auris. 2017. The variability of syllable patterns in Tashlhiyt Berber and Polish. *Journal of Phonetics* 64. 127–144.

Hermes, Anne, Rachid Ridouane, Doris Mucke & Martine Grice. 2011. Kinematics of syllable structure in Tashlhiyt Berber: The case of vocalic and consonantal nuclei. *9th International Seminar on Speech production.*, 1–6.

Hu, Fang. 2016. Tones are not abstract autosegmentals. *Speech Prosody*, 302–306.

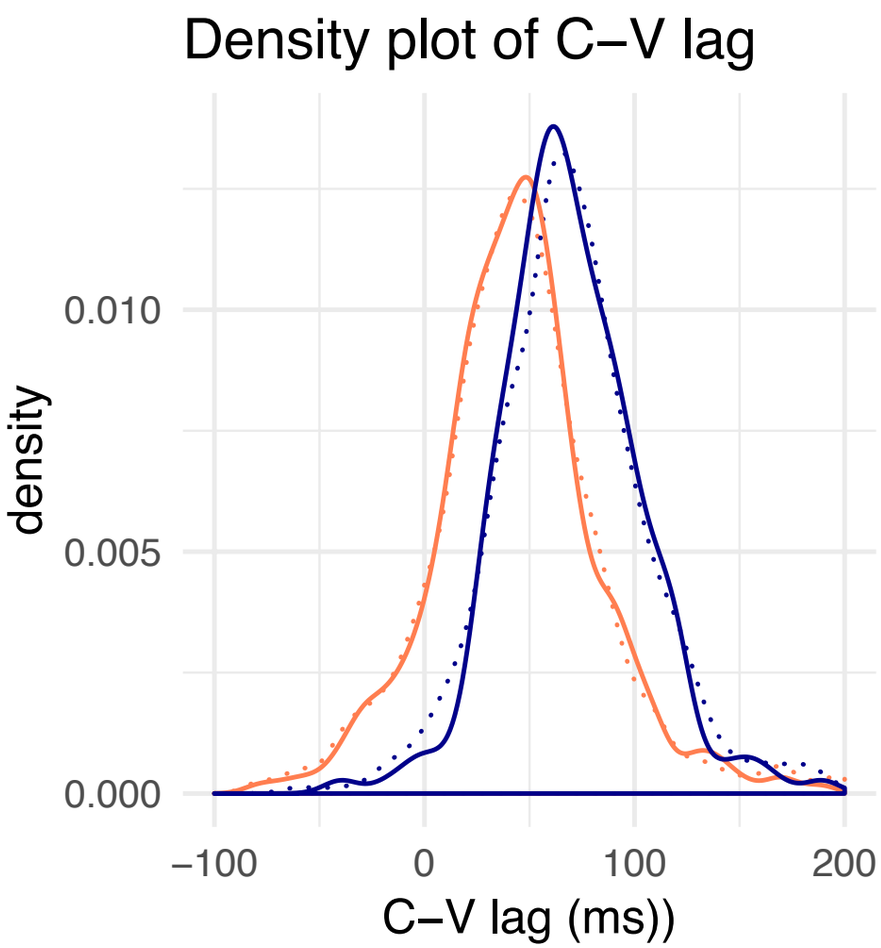
Karlin, Robin. 2014. The articulatory TBU: gestural coordination of tone in Thai. *Cornell Working Papers in Linguistics.*

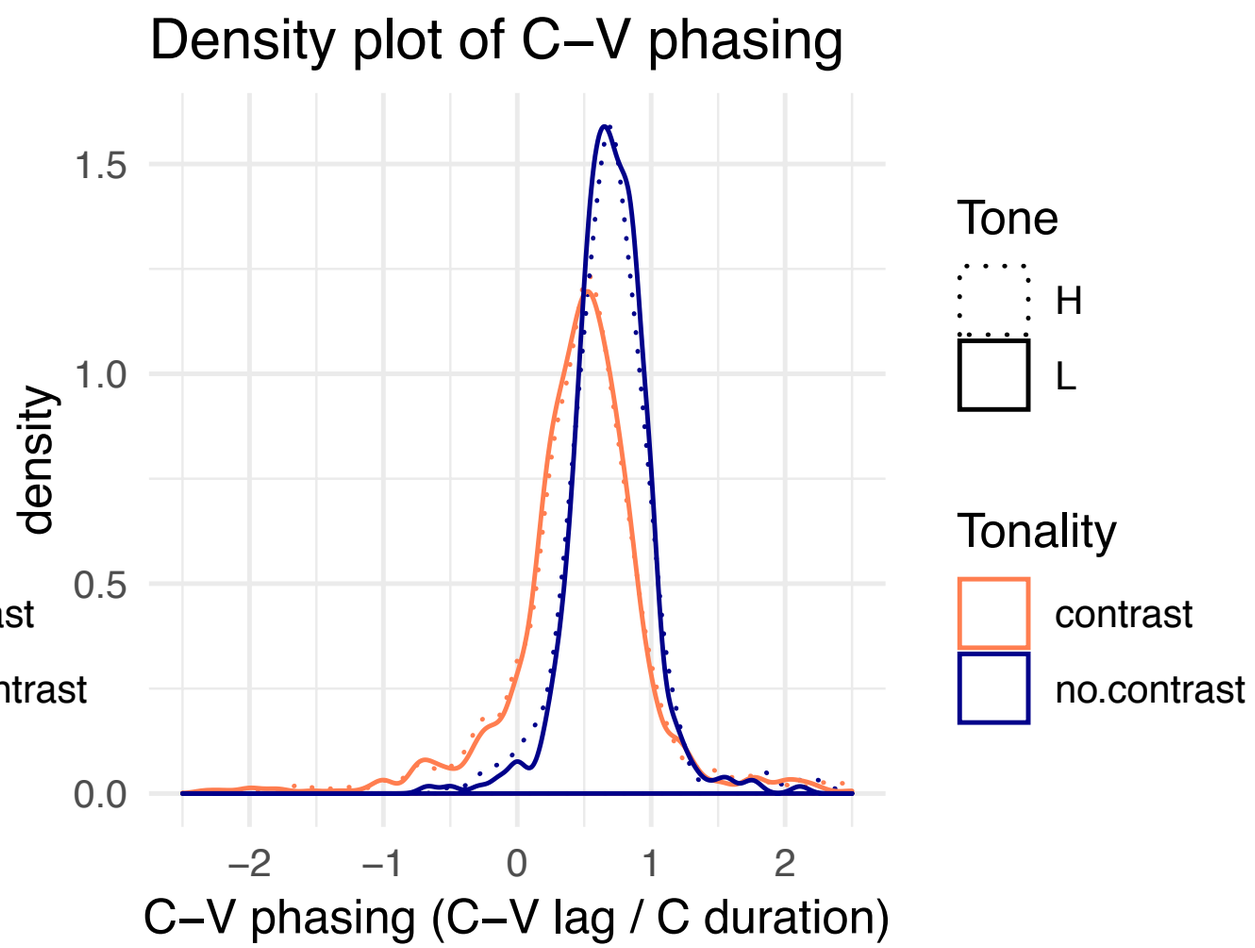
# References II

Katsika, Argyro, Jelena Krivokapić, Christine Mooshammer, Mark Tiede & Louis Goldstein. 2014. The coordination of boundary tones and its interaction with prominence. *Journal of Phonetics* 44. 62–82.
Marin, Stephania and Marianne Pouplier. 2010. Temporal Organization of Complex Onsets and Codas in American English: Testing the Predictions of a Gestural Coupling Model. *Motor Cotrol* vol. 14. p. 380-407
Mücke, Doris, Hosung Nam, Anne Hermes & Louis Goldstein. 2012. Coupling of tone and constriction gestures in pitch accents. In Philip Hoole, Bombien, Lasse, Pouplier, Marianne, Mooshammer, Christine & Barbara Kühnert (eds.), *Consonant clusters and structural complexity* (Interface Explorations), vol. 26, 205. De Gruyter Mouton.
Nam, Hosung & Elliot Saltzman. 2003. A competitive, coupled oscillator model of syllable structure. *Proceedings of the 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. *Proceedings of the 17th International Congress of the Phonetic Sciences*, 1486–1489.
Saltzman, Elliot et al. 2008. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In *SP-2008*, 175-184.
Shaw, Jason, Adamantios I. Gafos, Philip Hoole & Chakir Zeroual. 2009. Syllabification in Moroccan Arabic: evidence from patterns of temporal stability in articulation. *Phonology* 26(1). 187–215.
Tournadre, Nicolas & Sangda Dorje. 2003. Manual of standard Tibetan. Snow Lion Publications.

# C-V lag by tone

- No effect of tone on C-V lag



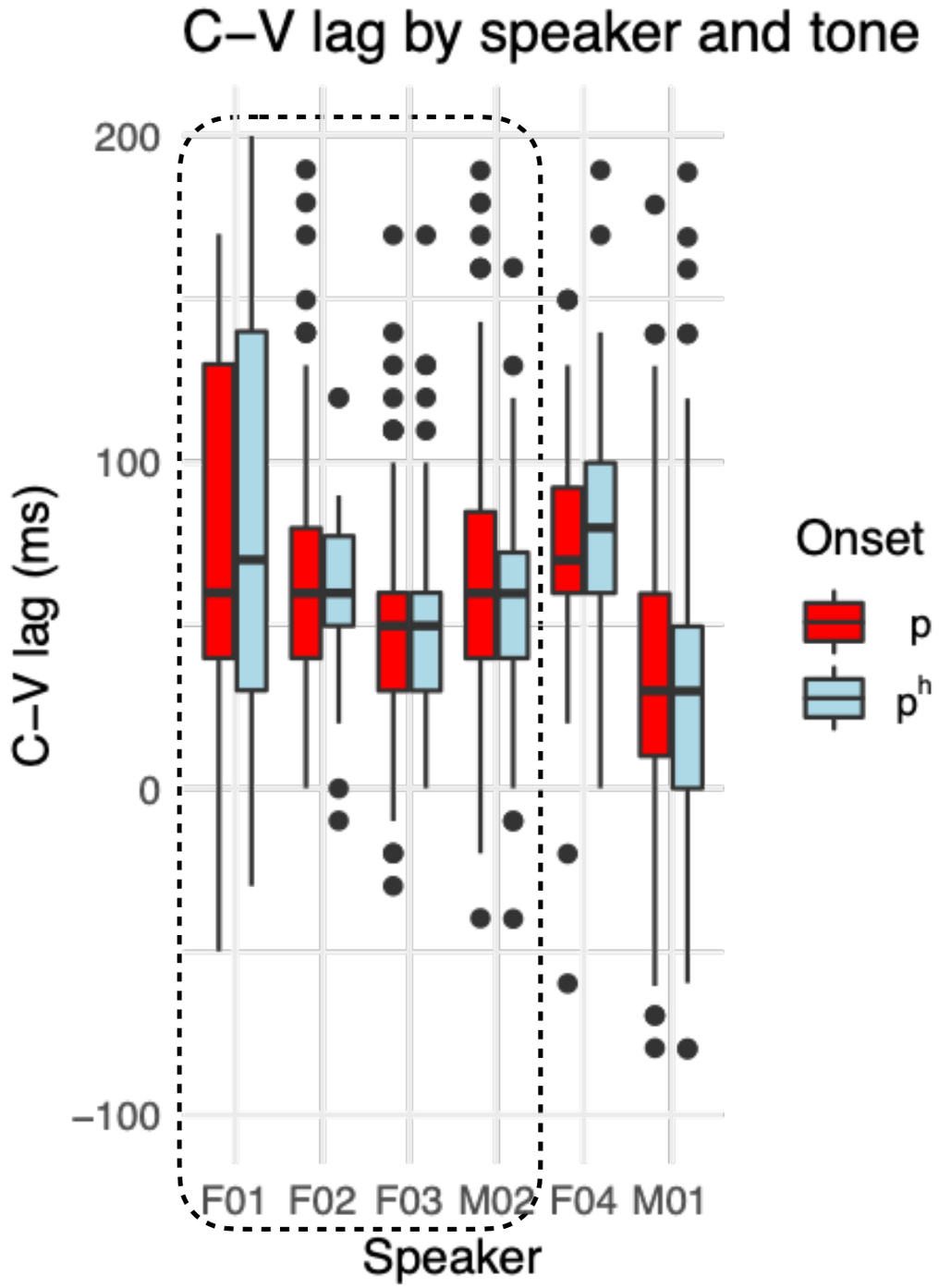



# Results: C-V timing

- C-V lag not signficantly different by aspiration either (/pV/ vs. /$p^h$V/)

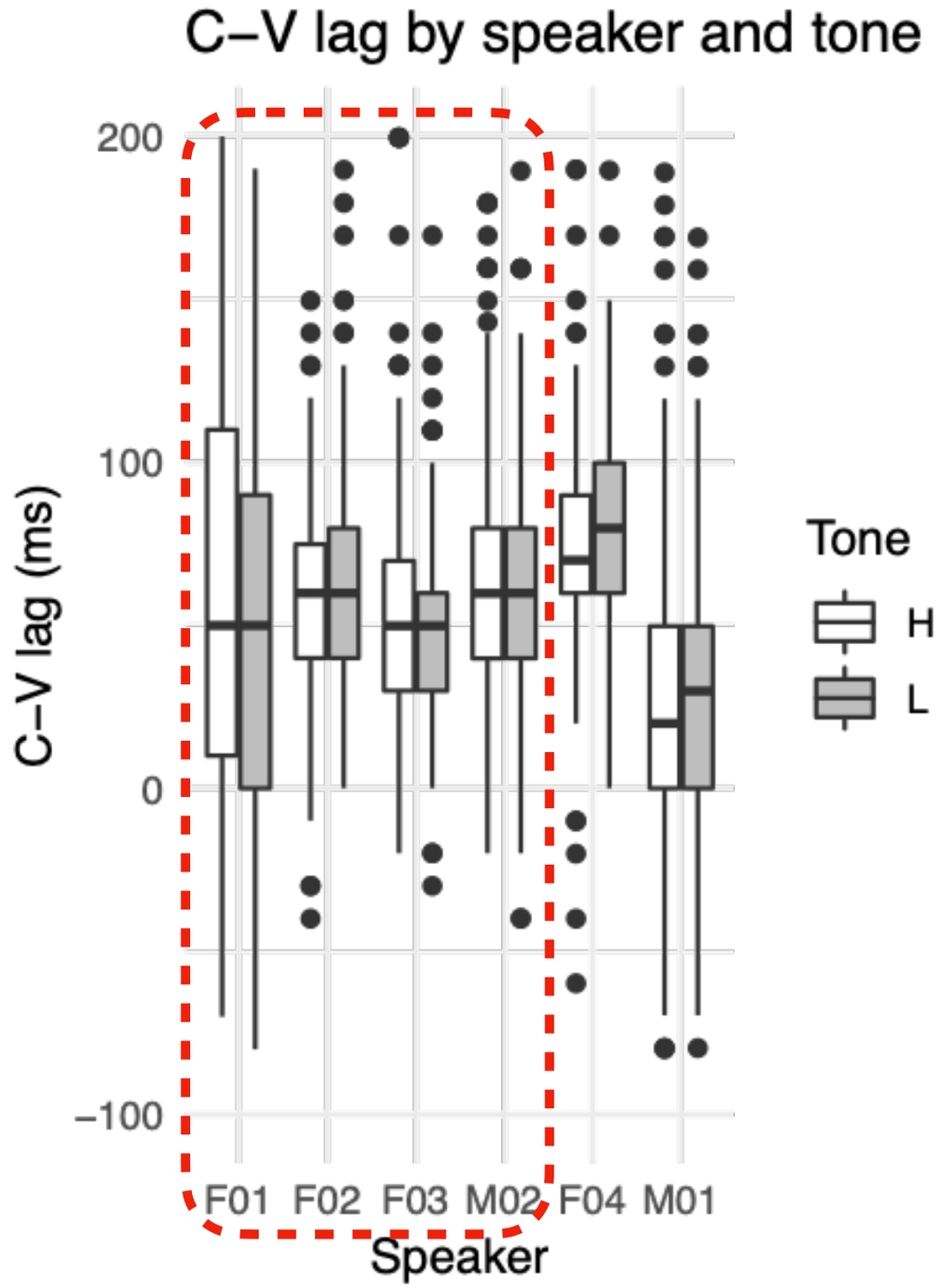
- LMM: random effects of speaker, word; fixed effect of tone contrast

    - model not improved by adding effect of onset

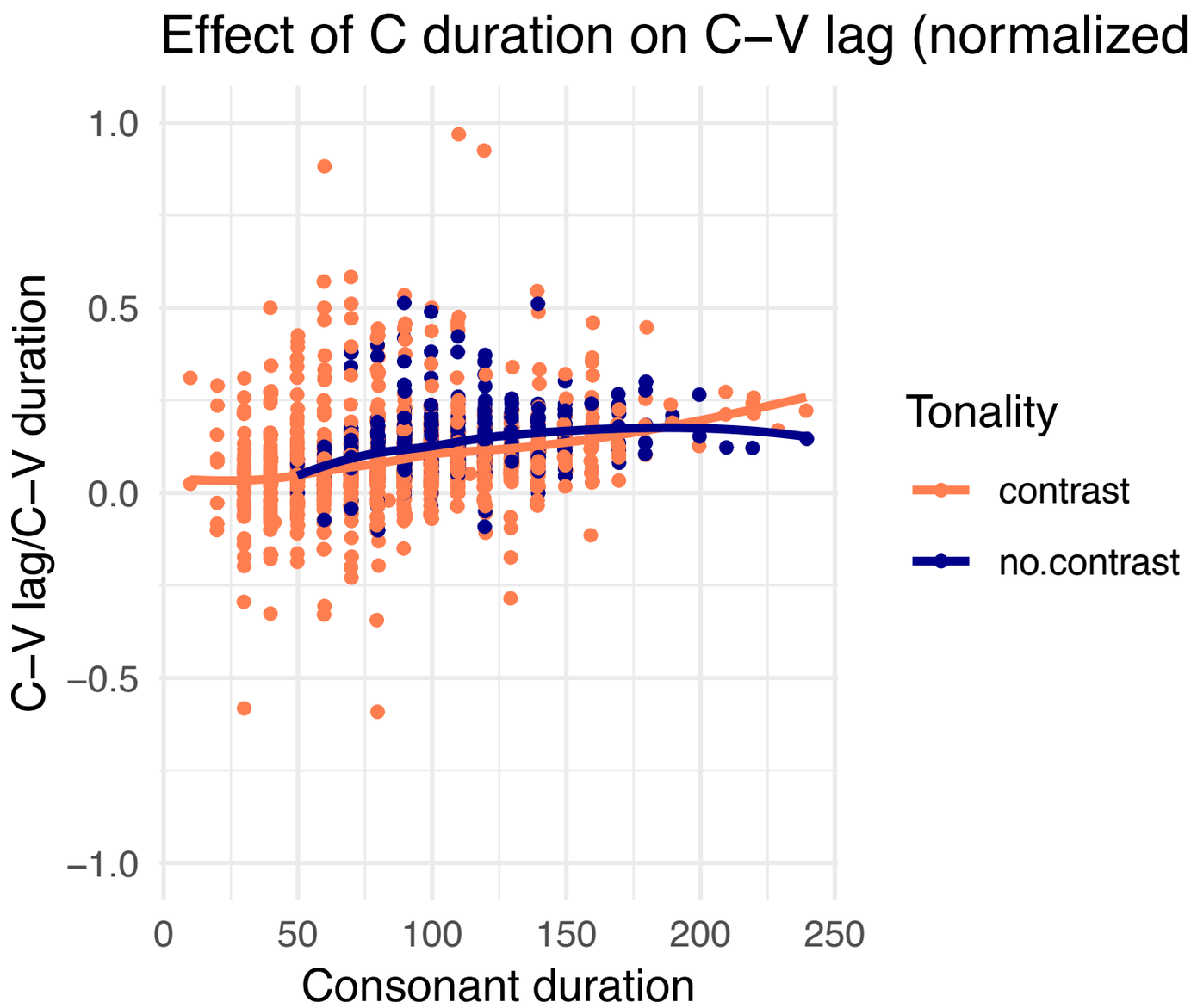


# Results: C-V timing

- C-V lag not significantly different across tones (/mV/)



# Results: C-V Phasing / Cdur



# Tibetan

- Tonal and non-tonal dialects
- Tonal dialects:
    - One tone per word
    - Two tone type: high-level and low-rising