Critically-damped oscillators and General Tau Theory exhibit similar error across speakers with different vocal tract dimensions

Christopher Geissler, Boston College
Jyothiraditya Nellakra, University of Southern California
geissle@bc.edu, nellakra@usc.edu

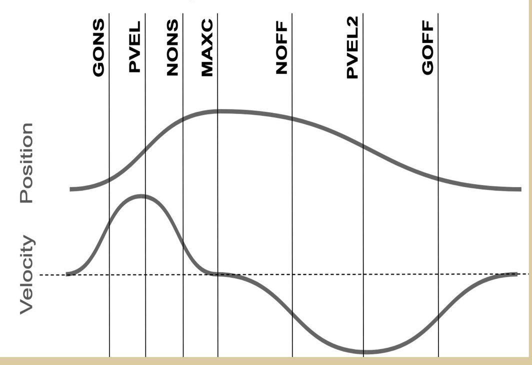
Models

- **How do we describe the shape of an articulatory gesture?**
  - Theories of control & planning → Equations with known parameters
- "CDO": Critically-Damped Oscillators
  - Mass-spring equation: $\ddot{x} + b\dot{x} + kx = 0$

$$x(t) = D\left(e^{-\omega_0 t} + \omega_0 t e^{-\omega_0 t}\right)$$

- "Tau": General Tau Theory
  - Gap-closure equation

$$x(t) = D\left(1 - \frac{t^2}{T^2}\right)^{\frac{1}{\kappa_{x,G}}}$$



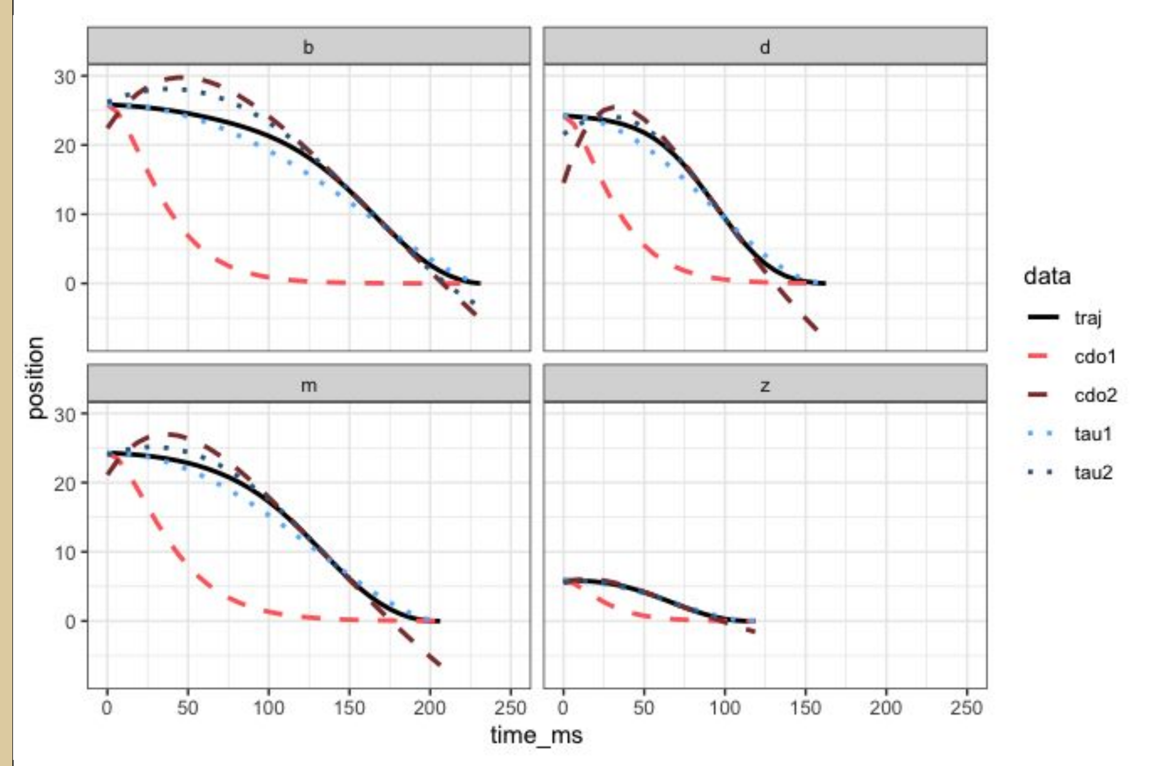
Methods

- X-Ray Microbeam Dataset
  - Closure of C in [əCɑ]: [p b m f v t d n s z ʃ ʒ]
  - Est. vocal tract length (VTL) from [ə] F3 & F4
- Unpack model equations to determine params
  - 2 additional parameters: *x*-offset and *t*-offset
  - CDO1 and Tau1: set parameters using GONS, PVEL, NONS. (5% vel. points + PVEL)
  - CDO2 and Tau2: set parameters using PVEL and first 20% velocity point.
  - CDO3 and Tau3: regress parameters to minimise error. (Theoretical best-fit)
- Error is mean Euclidean distance of modelled trajectory from experimental trajectory.
- Modelling and analysis was done using <High level summary of techniques and tools, details in next section>

Analysis & Results

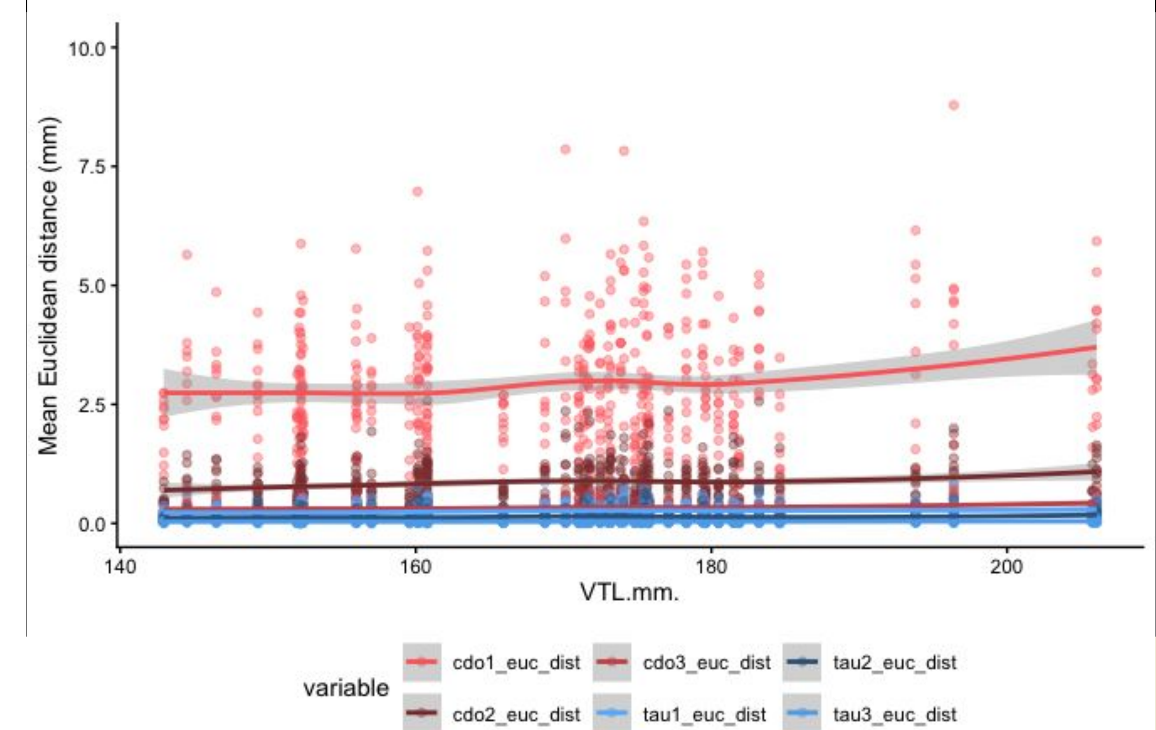
Sample fits

Sample simulations



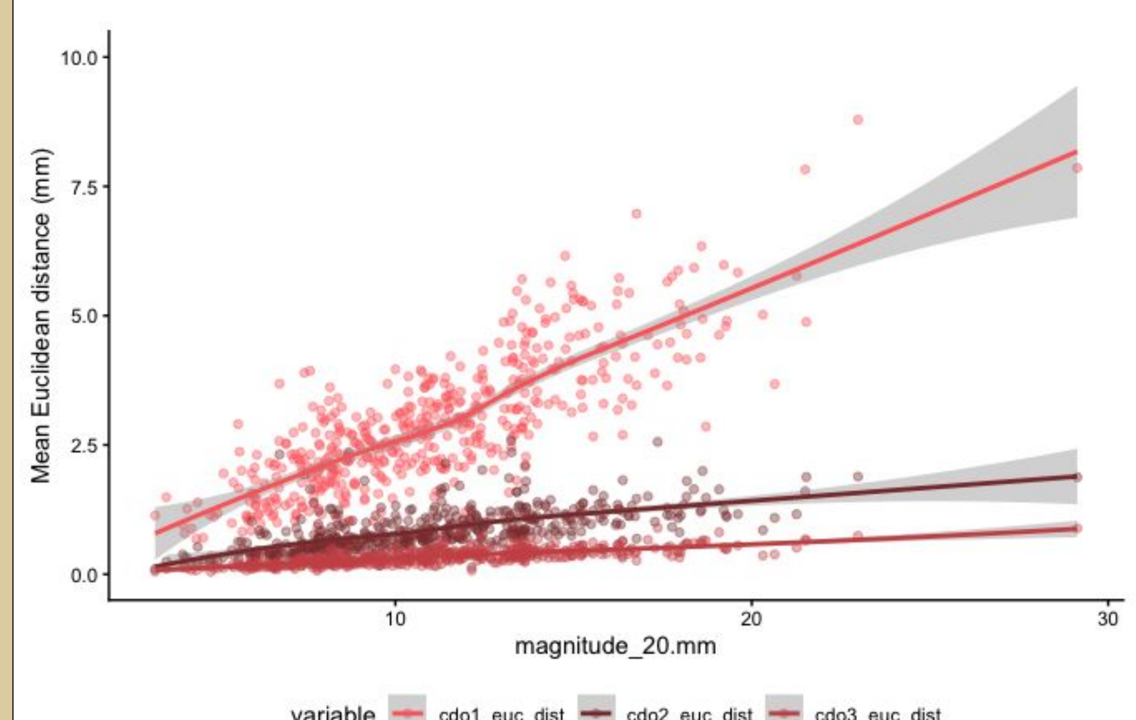
Vocal tract results

No effect of VTL/sex/height, if use magnitude
→ oral cavity height matters more than VTL



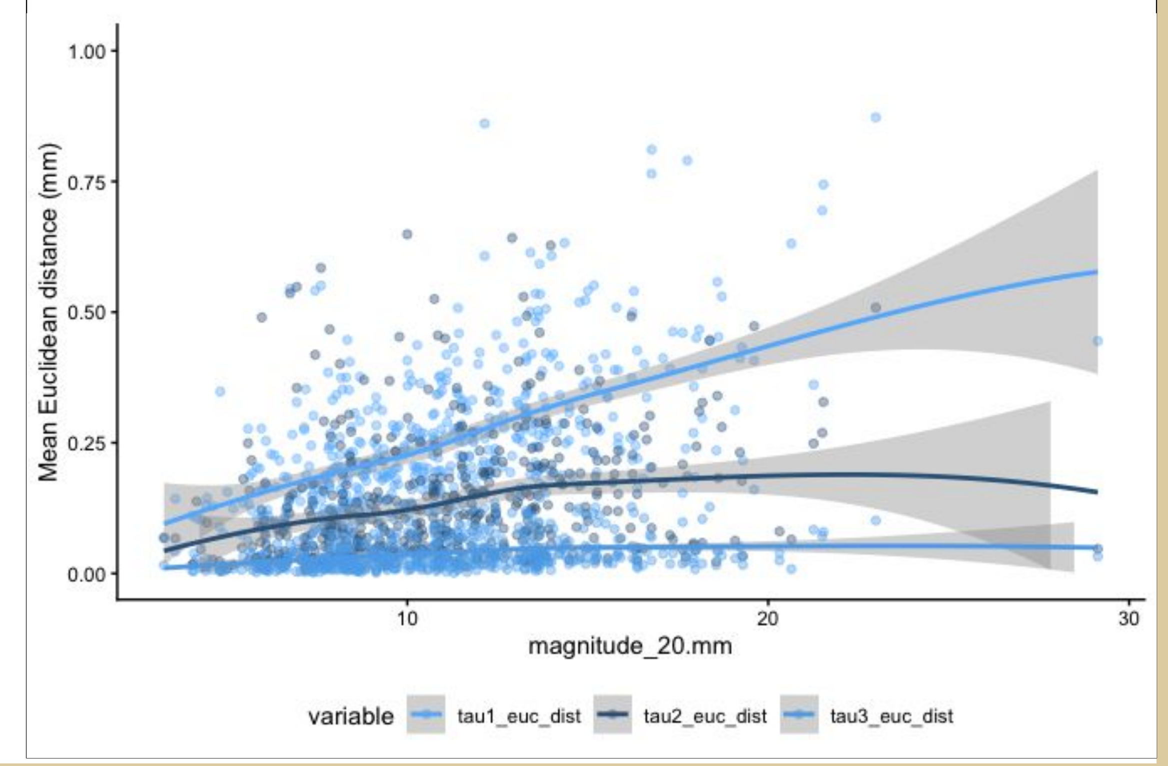
CDO results

CDO distance scales with magnitude
error ~ magnitude + manner + place



Tau results

Tau fits better than CDO (note y-axis scaling)
error ~ magnitude + manner + duration

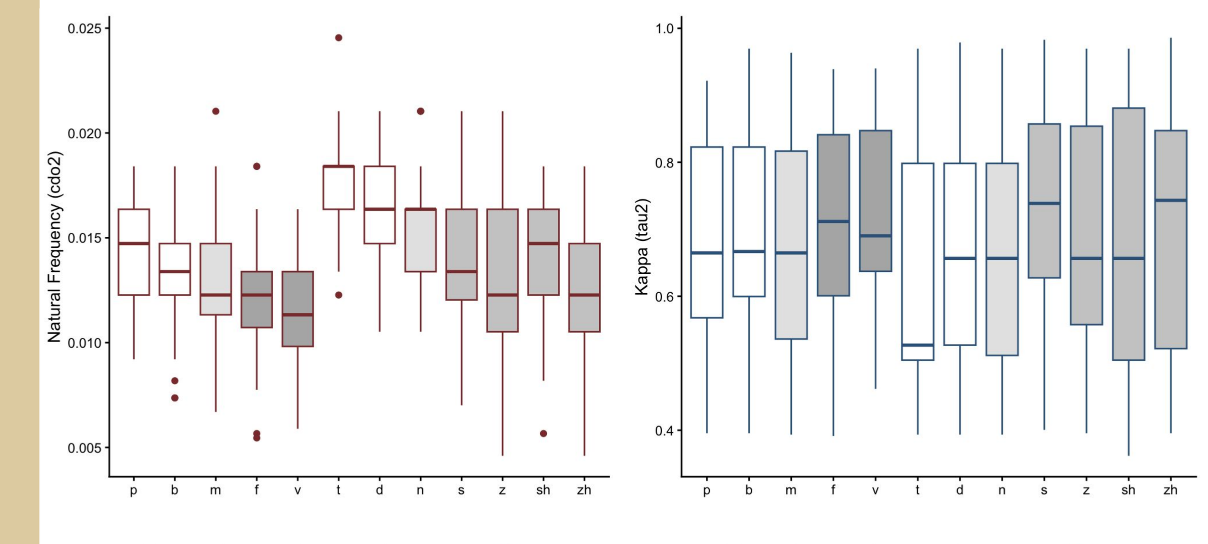


Unpacking Models

- Regressing parameters for best-fit cannot account for implicit statements of models:
  a. Acceleration at PVEL should be 0. → $\ddot{x}(t_{PVEL}) = 0$
  b. Velocity at GONS and NONS should be ≈ 0. → $\dot{x}(t_{GONS}) \approx \dot{x}(t_{NONS}) \approx 0$
  c. The velocity at the 20% points should be ⅕ of PVEL. → $\dot{x}(t_{20\%point}) = 0.2\dot{x}(t_{PVEL})$
- (a & b) → CDO1, Tau1. (a & c) → CDO2, Tau2
- *x*-offset and *t*-offset are also parameters.
  - Models treat start as *t* = 0 and end at *x* = 0.
  - Experimental data has arbitrary start times.
  - Goal might not be constriction degree *CD* = 0.
  - Conversion function requires *x* and *t* offset.
- Implications for a model of speech production:
  - Gesture execution requires two or three points (*x*, *vel*, *acc*, *t*). *x* is *CD* on a linear scale, *t* = 0 is arbitrary, depending on the gestural score.
  - Parameter setting for model *x*(*t*) is automatic.
  - Model *x*(*t*) function governs motor output.

Conclusions

- These models fit similarly across individuals
- New method of calculating model parameters
  - Below: Natural Frequency, Kappa
- Further work is needed to extend this analysis to models of gradient gestural activation over time



References

Elie, Benjamin, David N. Lee & Alice Turk. 2023. Modeling trajectories of human speech articulators using general Tau theory. *Speech Communication* 151. 24–38. https://doi.org/10.1016/j.specom.2023.04.004.

Lee, David N. 1998. Guiding Movement by Coupling Taus. *Ecological Psychology* 10(3/4). 221. https://doi.org/10.1080/10407413.1998.9652683.

Saltzman, Elliot L. & Kevin G. Munhall. 1989. A Dynamical Approach to Gestural Patterning in Speech Production. *Ecological Psychology* 1(4). 333–382. https://doi.org/10.1207/s15326969eco0104_2.

Westbury, John R., Greg Turner & J. Dembowski. 1994. *X-ray microbeam speech production database user's handbook, version 1.0*. University of Wisconsin Waisman Center.