Kinematic modeling of eyebrow raising and lowering gestures in ASL

Introduction. Low-order dynamical systems with critical damping have proven successful in accounting for the spatiotemporal properties of many target-directed linguistic gestures (e.g., lip protrusion and tongue tip raising; Saltzmann & Munhall, 1989). In this study, we investigated whether the same models can quantitatively characterize the kinematics of nonmanual movements in signed languages, focusing on eyebrow raising and lowering gestures in American Sign Language (ASL). Previous research has identified many lexical and grammatical functions of these gestures (e.g., Wilbur, 2021); here we study their internal formal properties.

Data. The American Sign Language Linguistic Research Project (ASLLRP) provides videos and expert annotations of isolated utterances produced by several native ASL signers (Neidle et al., 2022). We selected two signers (here S1 and S2) for which substantial data was available and analyzed all intervals annotated with **(further) raised / (further) lowered eyebrows** (S1: N = 857; S2: N = 531). The videos show a signer's face in close-up (‘camera 2’) at approximately 30 fps. Nonmanual gestures in portions of the same data have been analyzed previously (Liu et al., 2014), but their detailed kinematic profiles have not been studied.

Facial landmarks were tracked using the vision model of Bulat & Tzimiropoulos (2017) (https://github.com/1adrianb/face-alignment), which was designed to acccurately recover 3D locations from (static) 2D images regardless of head tilt, rotation, etc. For each video frame, the tracked landmarks were transformed to a canonical scale and front-facing coordinate system (Procrustes algorithm). Eyebrow height was calculated by averaging the vertical coordinates of the landmarks for each brow and subtracting the average height of the lower eye landmarks below it. We resegmented the expert ASLLRP annotations so that transitions into and out of raising/lowering gestures had values close to the average of neutral frames (neither raised nor lowered), which we set to zero. Figures below show the eyebrow trajectories after z-scoring and winsorizing at ± 3 sd within signer and then averaging over the two brows. While most raising gestures have a distinct rise–hold–fall profile (and lowering gestures show the inverse), individual movements clearly vary in the rate (or ‘stiffness’) of rise/fall and height of the target.

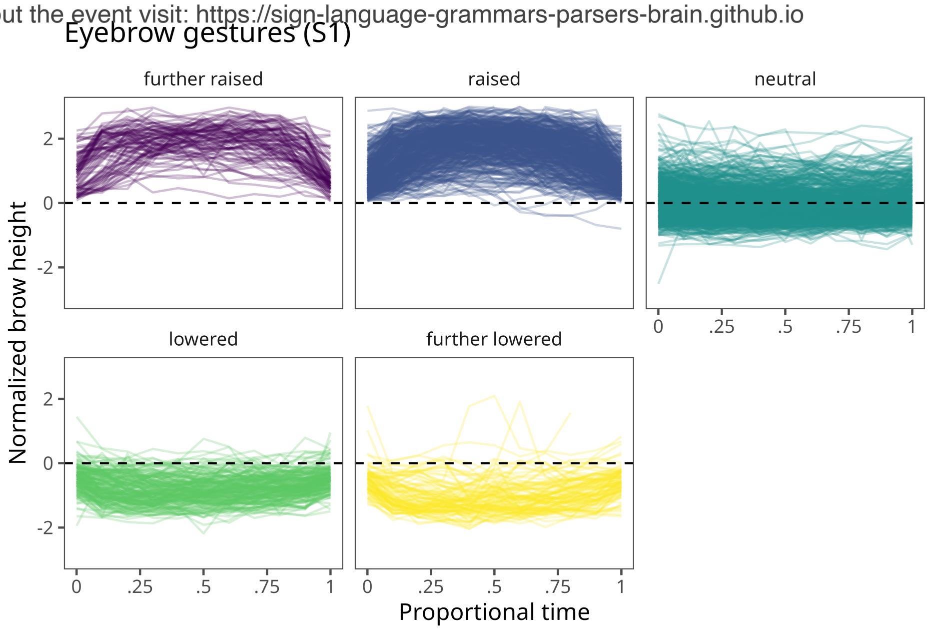
Models. We evaluated three models on their ability to provide accurate low-dimensional representations of these nonmanual trajectories. The first model (TD2) combines two gestural components, one for the initial rise/fall and another for the final fall/rise, with the relative activation of the two components governed by a sigmoidal function of time. (Notation: $\ddot{x}$ acceleration, $\dot{x}$ velocity, x height, κ stiffness, τ target height. Greek letters are free parameters.)

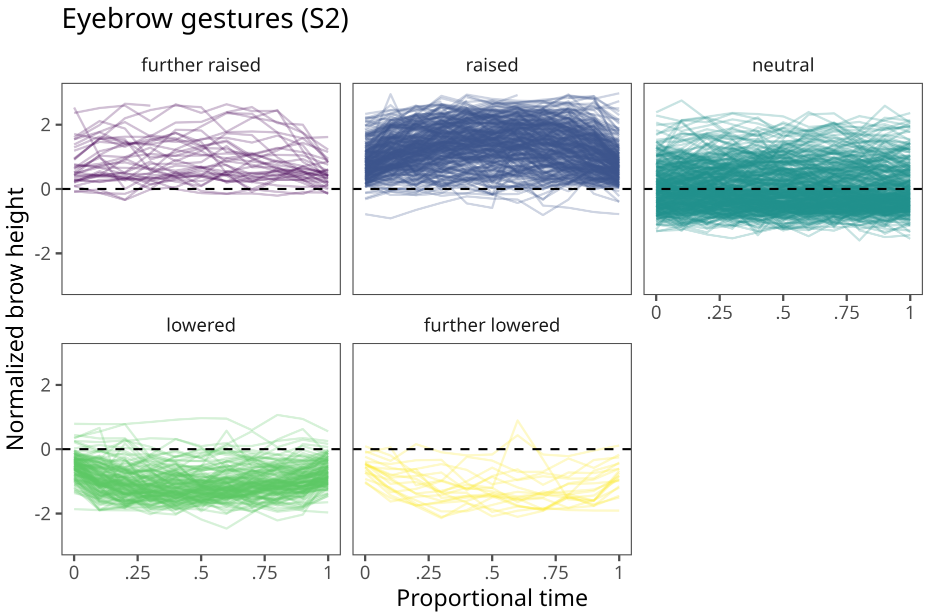
$$\ddot{x} + \sqrt{4k(t)}\dot{x} + k(t)(x - T(t))$$
$$\text{a(t)} = 1 \text{ - sigmoid(t; } \mu, \sigma)$$
$$\text{k(t)} = \text{a(t)*}\kappa_{begin} + (1 - a(t)) * \kappa_{end}$$
$$\text{T(t) *} = \text{a(t)*}\tau_{begin} + (1 - a(t)) * \tau_{end}$$

The second model (TD3) is identical except for the addition of a cubic term motivated by the finding that, in speech, peak velocity occurs near the middle of displacement within a gesture (Sorensen Gafos, 2016). The third model (Tau) has been argued to capture velocity/displacement relationship with one fewer parameter than TD3; however, because it accounts for movements toward targets but not prolonged holds, it also needed an additional parameter.

Results. Models were fit to individual gestures and evaluated by pseudo-R^2 ($1 -$ SS residual / SS total). All three models were highly accurate for raising gestures (TD2: *mean* = 0.81 [.025-*quantile* = 0.25, .975-*quantile* = 0.86]; TD3: 0.83 [0.33, 0.88]; Tau: 0.81 [0.25, 0.86]), somewhat below an unrestricted fifth-order polynomial baseline (0.90 [0.51 0.91]). Performance on lowering gestures was worse for all models (grand mean = 0.61) including the baseline (0.69). These results suggest that the simplest, most restrictive dynamical model (TD2) can provide a competitive account of eyebrow raising gesture kinematics, and motivate further investigation of the measurement and kinematics of eyebrow lowering.

This abstract has been accepted for presentation at the Interdisciplinary Workshop "Sign Language Grammars, Parsing Models, & the Brain", 6-7 November 2025, Max Planck Institute for Human Cognitive & Brain Sciences, Leipzig, Germany. For further information about the event visit: https://sign-language-grammars-parsers-brain.github.io





References. Bulat, A., Tzimiropoulos, G. (2017). How far are we from solving the 2D & 3D face alignment problem? (and a dataset of 230,000 3D facial landmarks). In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1021-1030). • Elie, B., Lee, D. N., Turk, A. (2023). Modeling trajectories of human speech articulators using general Tau theory. *Speech Communication*, 151, 24-38. • Liu, J., Liu, B., Zhang, S., Yang, F., Yang, P., Metaxas, D. N., Neidle, C. (2014). Non-manual grammatical marker recognition based on multi-scale, spatio-temporal analysis of head pose and facial expressions. Image and Vision Computing, 32(10), 671-681. • Neidle, C., Opoku, A., Metaxas, D. (2022). ASL Video Corpora & Sign Bank: Resources available through the American Sign Language Linguistic Research Project (ASLLRP). arXiv preprint arXiv:2201.07899. • Saltzman, E. L., Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, 1(4), 333-382. • Sorensen, T., Gafos, A. (2016). The gesture as an autonomous nonlinear dynamical system. *Ecological Psychology*, 28(4), 188-215. • Wilbur, R. B. (2021). Non-manual markers: Theoretical and experimental perspectives. In *The Routledge Handbook of Theoretical and Experimental Sign Language Research* (pp. 530-565). Routledge.