Using Machine Learning to Build an Inventory of Sign Components

Lee Kezar | University of Southern California **Zed Sehyr** | Chapman University **Jesse Thomason** | University of Southern California

Can machine learning (ML) models identify the **discrete linguistic units** of signs by analyzing articulatory patterns in **isolated sign productions**?

Background

Articulatory variation in signing exceeds what is necessary for recognition, similar to speech. However, the relationship between variation and the hypothesized abstract components of signs (e.g., phonemes) remains an open question.

Research Question

How do ML-generated components correspond to the phonological units of signs?

Method

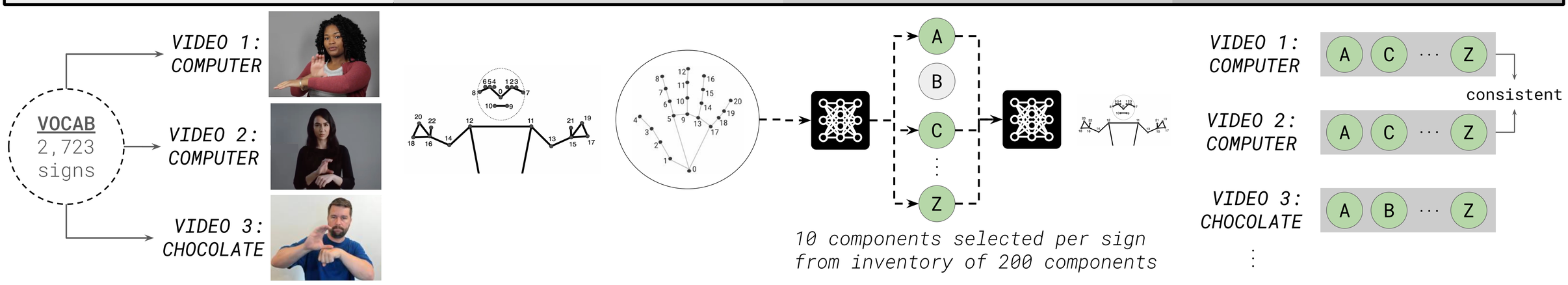
Develop a self-supervised ML approach to decompose signs into **discrete units**, and investigate their correspondence to phonological structures. The units are labels, learned by the model, that correspond to specific visual and spatial patterns.

STEP 1
Isolated sign videos (n=175k) in ASL are converted into frame sequences

STEP 2
Position of signers' joints (75 X/Y pts) extracted with MediaPipe [1]

STEP 3
Model identifies the components in each sign and reconstructs the pose

STEP 4
We evaluate the consistency and accuracy of the identified components



Key Finding Components generated by our model correspond to ASL phonemes: they accurately (72%) predict the sign identity and correlate with hypothesized phonemes in ASL-LEX.

Datasets Used

Sem-Lex Benchmark

- **91k isolated sign videos** from deaf signers (n=41) of American SL [2]
- Participants report early exposure to ASL.
- Each video is manually labeled with ASL-LEX sign IDs (n=2,723) [3]

ASL Citizen

- **84k isolated sign videos** from deaf signers (n=52) of American SL. [4]
- Each video is manually labeled with ASL-LEX sign IDs (n=2,723)

The identified components can discriminate signs

→ Model ***consistently identifies*** similar components from multiple productions by different signers

→ The ***discriminative capacity*** of the model suggests that phonological contrasts in sign languages can be identified computationally.

Some components correspond to ASL phonological features

→ We compared the ML-generated components to 191 ***phonological features*** listed in ASL-LEX [3].

→ A separate ML model was able ***predict*** phonological features (e.g., handshape or movement type) from the learned components → these may have some linguistic relevance.

Limitations & Next Steps

→ The identified components are ***not immediately interpretable***, and ~28% of phonologically distinct signs are ***incorrectly labeled*** with the components of another sign.

→ ***Planned study***: Verify that the model can reconstruct signs from human (signer) judgements.

→ Can the model capture hierarchical relationships, such as handshape decomposed into ***finger flexion*** or ***spread***?

→ Can this model generalize ***across sign languages*** and reveal shared structures with speech?

USC University of Southern California
Grounding Language in Actions, Multimodal Observations, and Robotics Lab (GLAMOR)
< www.glamor.rocks >
CHAPMAN UNIVERSITY
COGNITION LANGUAGE & PLASTICITY LAB
< www.claplab.org >

[1] Lugaresi, C., [...] Uboweja, E., Hays, M., Zhang, F., Chang, C., Yong M., Lee, J., Chang, W., Hua, W., Georg, M. and Grundmann, M. (2023). "MediaPipe: A Framework for Perceiving and Processing Reality". In: IEEE Computer Vision and Pattern Recognition (CVPR)
[2] Kezar, L., Pontecorvo, E., Daniels, A., Baer, C., Ferster, R., Berger, L., Thomason, J., Sehyr, Z., and Caselli, N. (2023). "The Sem-Lex Benchmark: Modeling ASL Signs and Their Phonemes". In: ACM SIGACCESS Conference on Computers and Accessibility (ASSETS)
[3] Sevcikova Sehyr, Z., Caselli, N., Cohen-Goldberg, A. M., & Emmorey, K. (2021). "The ASL-LEX 2.0 Project: A database of lexical and phonological properties for 2,723 signs in American Sign Language". In: Journal of Deaf Studies and Deaf Education.
[4] Desai, A. Berger, L., Minakov F., Milan, V., Singh, C. Pumphrey, K., Ladner, R., Daumé, H., Lu, A. Caselli, N. Bragg, N. "ASL Citizen: A Community-Sourced Dataset for Advancing Isolated Sign Language Recognition" In: Advances in Neural Information Processing Systems (NeurIPS).