Measuring changes in Nicaraguan signing using skeletal modeling with analog 2D video sources

Ann Senghas[1], Aline Normoyle[2], Naomi Caselli[3], Corrine Occhino[4], and Matthew Dye[5]
[1]Barnard College, [2]Bryn Mawr College, [3]Boston University, [4]University of Texas at Austin, [5]Rochester Institute of Technology

Introduction

How do characteristics of the human body shape sign languages?

We are developing tools to measure quantitative aspects of signing to determine whether:

- these aspects change over time
- they represent adaptations to articulatory or perceptual constraints
- change occurs with repeated use, over the lifespan of a signer, or in transmission to new learners.

- We are leveraging advances in computer vision[1] to fit a hierarchical 3D skeletal model to 2D data, constrained in size, mass, and proportion (Fig. 2).
- Then we extract from that skeletal model estimates of distance, velocity, and force. These are used to compute meaningful metrics such as articulatory effort, symmetry, and production in specified zones (Fig 3).

Fig 2: Pipeline for extraction of sign-language metrics from video.

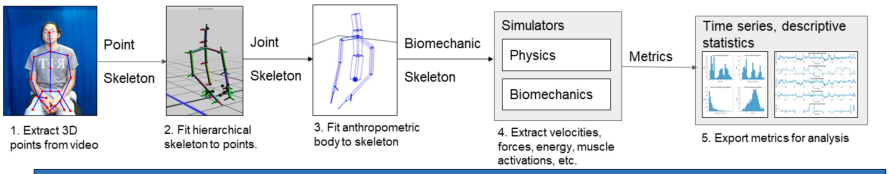
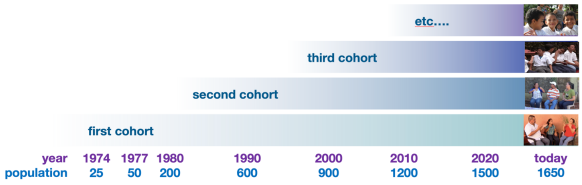


Fig 1: History of Lengua de Señas Nicaragüense

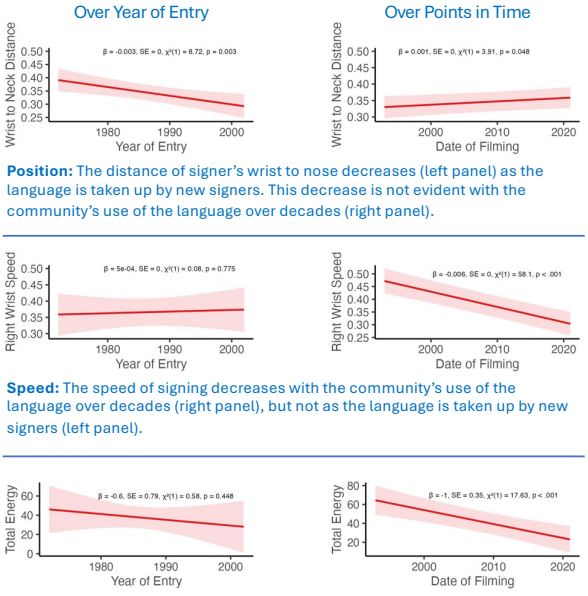


Results

For each video we generate frame-by-frame pose estimates, from which we compute:

- **Position** of signs, as distance from wrist to nose
- **Speed**, as velocity of movement of right wrist
- **Effort**, as kinetic energy required for sign production

Fig 5: Changes in signs over year of entry and point in time



Position: The distance of signer's wrist to nose decreases (left panel) as the language is taken up by new signers. This decrease is not evident with the community's use of the language over decades (right panel).

Speed: The speed of signing decreases with the community's use of the language over decades (right panel), but not as the language is taken up by new signers (left panel).

Effort: The effort expended in sign production decreases with the community's use of the language over decades (right panel), but not as the language is taken up by new signers (left panel).

Method

- Our data are drawn from an archive of Lengua de Señas Nicaragüense (LSN), which arose in the 1970s among deaf schoolchildren in Managua, Nicaragua. It has been transmitted to new learners since, and is the primary language of 1650 signers (Fig 1).
- The archive includes over 1000 hours of 2D video dating from 1990 to today.
- The present study analyzes 548 short narratives elicited using cartoons (*Koumal* and *Canary Row*) from 79 LSN signers, who all began signing at age 5 or younger.
- We compare signers across different years of entry into the community to discover the changes due to transmission, during acquisition upon initial contact in childhood.
- We compare samples of the language at different points in time to discover the changes due to years of experience with the language over signers' lifespans.

Fig 3: Heatmap of wrist positions.

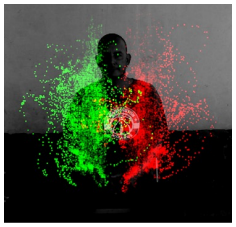
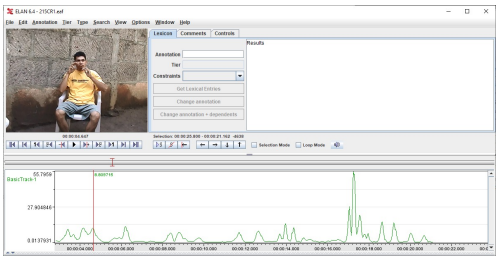


Fig 4: Kinetic Energy on a tier in ELAN[2]



Discussion

- Previous work across sign languages has suggested that perceptual pressures have shaped the properties of signs over time[3,4].
- New modeling work quantifying synchronic variability suggests that indeed, over time signs moved closer to the face[5] as a function of specific properties[6].
- In our diachronic, cross-generational approach, we tease apart the effects of **transmission over generations** (year of entry) and **use of a language** over years within a community of signers (date of filming).
- By combining pose estimation and skeletal modeling algorithms, we derived both perceptual and articulatory metrics.
- We found that over decades of transmission, new signers changed LSN, producing signs closer to the face
- We found that over decades of use, with increased experience within the signing community, LSN increased in articulatory efficiency, with a slower signing speed requiring less energy.

References

1. Artacho, B. & Savakis, A. (2020). Unipose: Unified human pose estimation in single images and videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7035–7044.
2. Crasborn, O., & Sloetjes, H. (2008). Proceedings of LREC 2008
3. Woll B. (1987). In Kyle JG, ed. Sign and school. pp12-34.
4. Siple P. Visual constraints for sign language communication. Sign Language Studies. 1978;19(95-110).
5. Flaherty, M., Sato, A., & Kirby, S. (2023). Cognition, 47 (4), 13277
6. Caselli, N., Occhino, C., Artacho, B., Savakis, A., & Dye, M. (2022). Cognition, 224, 105040.

Funding

US National Institutes of Health NIDCD grant R01 DC005407 to AS; US National Science Foundation BCS-1749384 to NC and BCS-1749376 to MD.