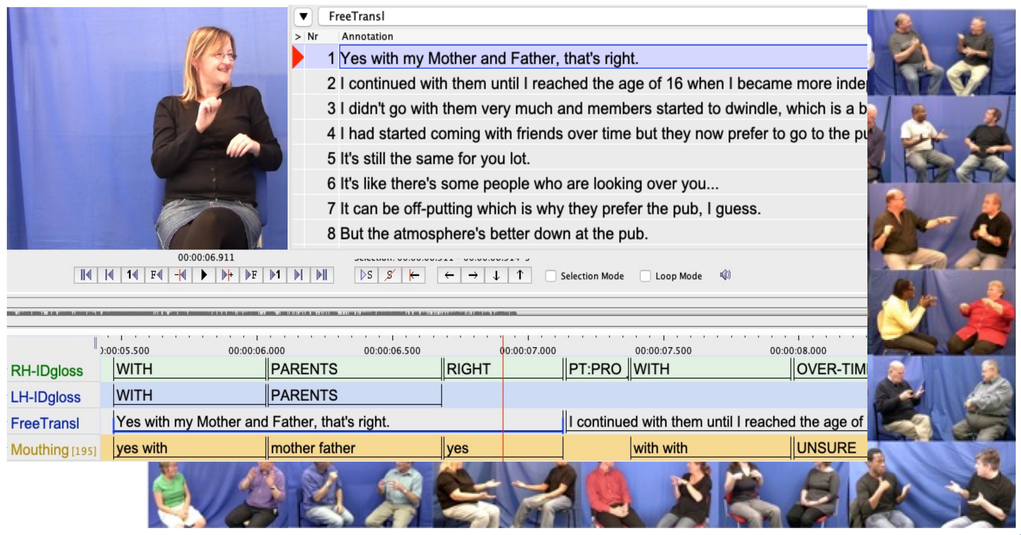
Segmentation of Signs for Research Purposes: Comparing Humans and Machines

Bencie Woll, Neil Fox, Kearsy Cormier
DCAL, University College London

DCAL
DEAFNESS COGNITION AND LANGUAGE RESEARCH CENTRE

BSL Corpus

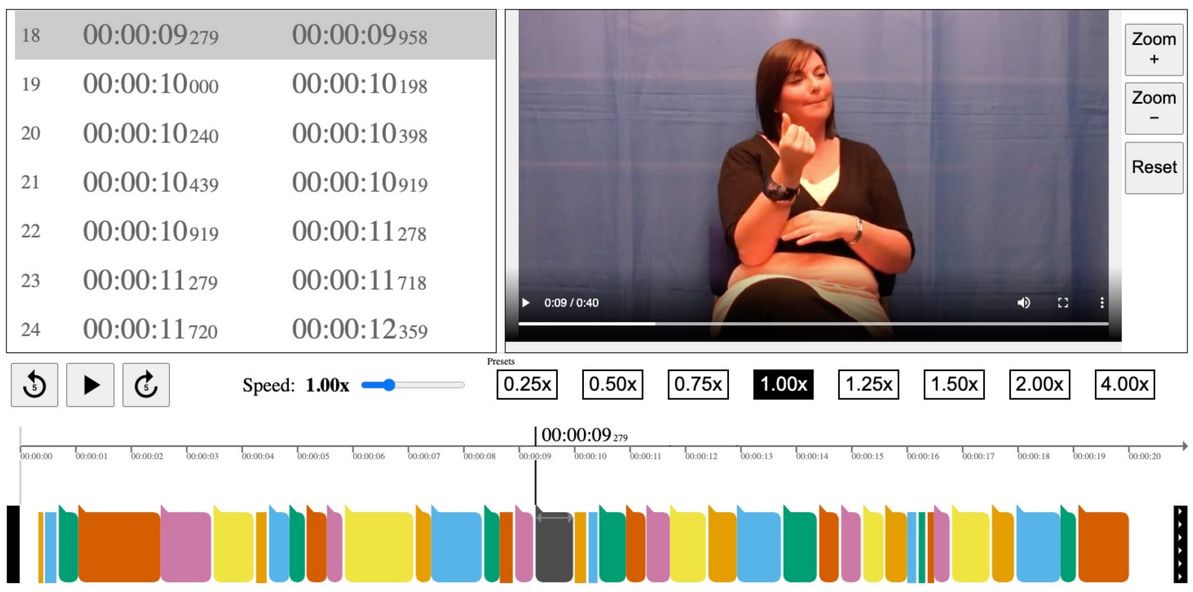
- Collection of ~125 hours of videos & metadata from 249 deaf signers from 8 UK regions: bslcorpusproject.org
- Videos online since 2011, initial annotations since 2014
- Annotation is SLOW – in 2022, 2/3 remains unsegmented



Sign Segmentation Tool: VIA-SLA

VIA-SLA via the Google Chrome browser

- www.robots.ox.ac.uk/~vgg/research/signsegmentation/
- Tech requirements at testing: videos < 1 minute, < 5MB
- Current file size limit is at 256MB



Segmentation by Humans

ID Glossing including segmentation

- 200 x real time for basic ID glossing of SL data
- 1 min of video: 3 to 4 hours for annotation (Crasborn 2015)

Segmentation alone

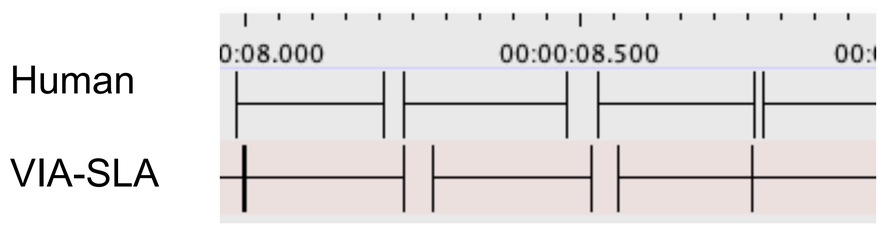
- 1 min of video: 1 hour by skilled annotator for segmentation (Mostowski et al. 2018)

Research Question

Can machine learning technology help speed up the segmentation process?

Methods

- Four x 1 min videos manually segmented by human (2nd author) and by VIA-SLA
- Annotation start and end time, and time needed for annotation, was recorded



- Machine segmentations with start/end within 100ms of human were coded as accurate
- Intelligibility of machine segmentations checked to ensure predicted segment contained a single sign

Preliminary Results

Video number	Duration (sec)	Time Manual Segmentation (sec)	Tool Predictions (sec)	No of Manual Segments	No of Predicted Segments	Predicted Segments Accepted	% Prediction Accuracy
1	14.5	480	21	24	29	20	83.3
2	40	1200	73	89	86	68	76.4
3	27	1020	59	42	39	31	73.8
4	19	660	26	31	35	24	77.4
Mean	25.1	840	44.8				77.7

- Use of VIA-SLA for segmentation took 5.3% of the time needed for manual segmentation
- Mean prediction accuracy of VIA-SLA was around 78%

Next steps

- One important next step: measuring how long it takes a human to correct the machine annotations (then compare human time with machine+correction time)
- Much more testing needed with more and longer videos and with videos in other sign languages.
- Nevertheless, VIA-SLA is promising (and much needed) tool to help speed up annotation
- If performance and reliability can continue to improve, such a tool will ultimately prove very useful for sign linguists.

DCAL: Deafness, Cognition & Language Research Centre
University College London
49 Gordon Square, London, WC1H 0PD
Email: dcal@ucl.ac.uk Website: http://www.dcal.ucl.ac.uk

EPSRC
Engineering and Physical Sciences Research Council
UCL