Capturing Methodology for Generating Synthetic and 3D Training Data in Catalan Sign Language (LSC): The Case of Verbal Agreement

Universitat Pompeu Fabra Barcelona

Gemma Barberà, Inés Broto Clemente, Xavier Vinaixa Roselló, Roger Cassany Viladomat

signarIA

Sørensen AI for media, culture and the arts

1. Introduction and goal: Overcoming data scarcity

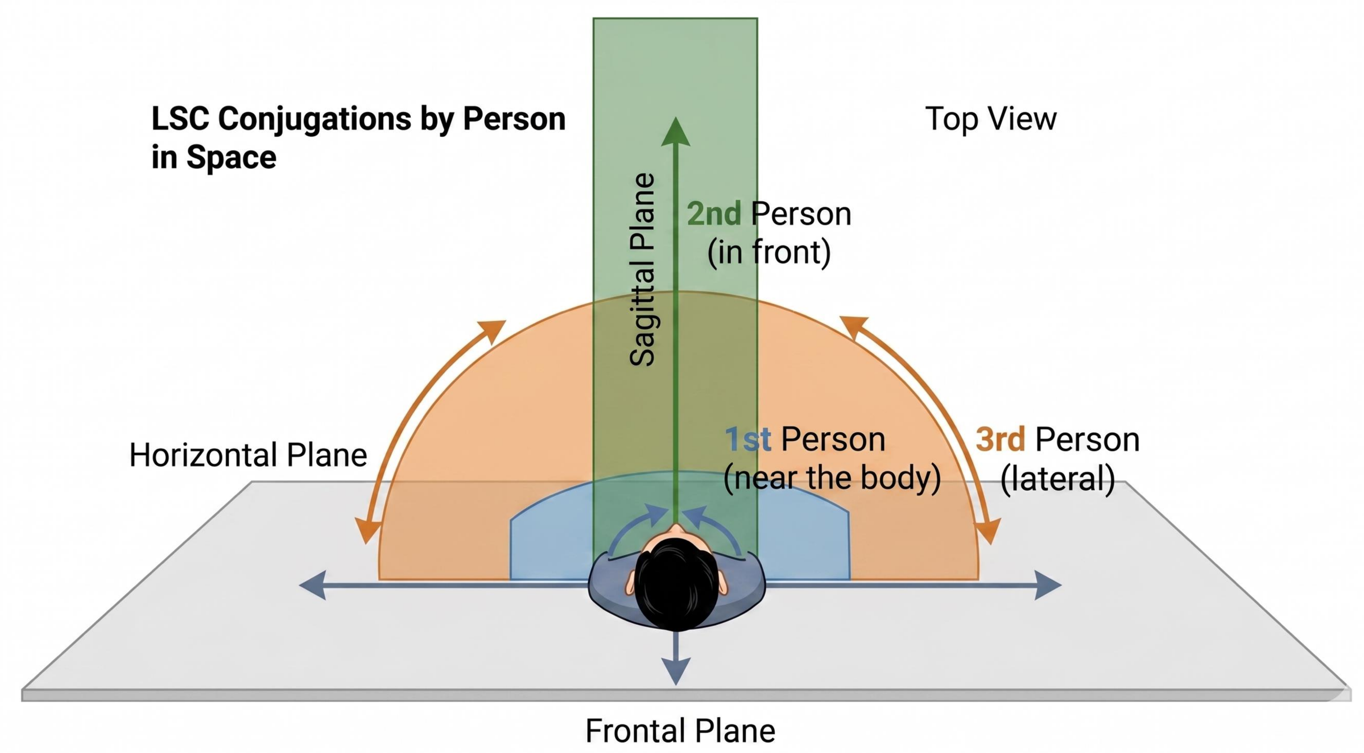
The problem: Sign Language Recognition systems require massive amounts of data, which is unfeasible to record manually for low-resource languages like Catalan Sign Language (LSC) [1].

Goal: To demonstrate that it is possible to synthetically expand a dataset by focusing on verb agreement, a domain known for its morphological and articulatory complexity.

2. Theoretical framework: Verb agreement and signing space

Agreeing verbs: Marking reference to arguments by linking the start and end points of a movement to spatial referential locations.

Spatial planes: Agreement is expressed across three structured spatial planes.



Gloss-level annotated dataset: Augment a gloss-level annotated dataset by leveraging the linguistic rules that govern sign–gloss transcriptions.

3. Corpus design: Calibration dataset

Complex repertoire of gramatical constructions including different features.

Person-marked forms of TELL:

Gloss	Translation
TELL	(infinitive)
1TELL2	'I tell you.'
2TELL1	'You tell me.'
1TELL3_a	'I tell him.'
3_aTELL1	'He tells me.'
1TELL3_b	'I tell her.'
3_bTELL1	'She tells me.'
2TELL3_a	'You tell him.'
3_aTELL2	'He tells you.'
2TELL3_b	'You tell her.'
3_bTELL2	'She tells you.'
3_aTELL3_b	'He tells her.'
3_bTELL3_a	'She tells him.'

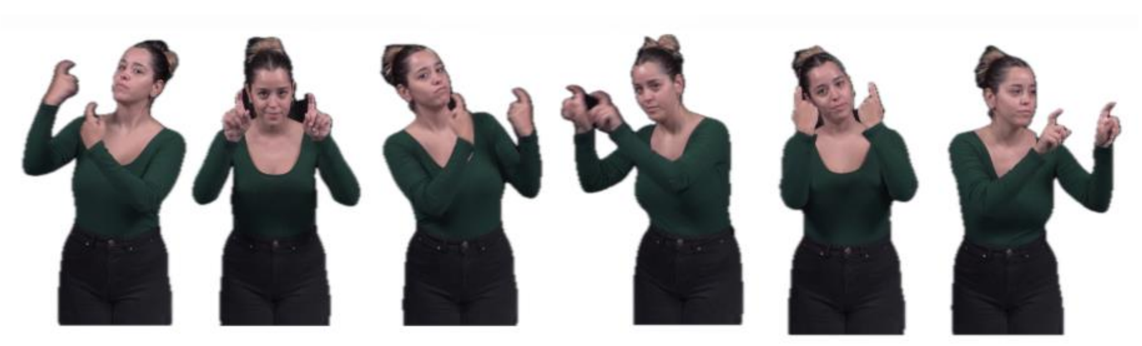
High/indefinite third-person subject forms of TELL [2]:

Gloss	Translation
3_{ahi}TELL1	'They/Someone tells me.'
3_{bhi}TELL1	'They/Someone tells me.'
3_{ahi}TELL2	'They/Someone tells you.'
3_{bhi}TELL2	'They/Someone tells you.'
3_{ahi}TELL3_b	'They/Someone tells her.'
3_{bhi}TELL3_a	'They/Someone tells him.'

Signer's profile and amount of data:

	Gender	Age group	Recordings
TG	M	51-80	949
BF	M	31-50	952
CW	F	18-30	952
TOTAL			**2853**

Various agreement forms of the verb ADVICE:



4. Multiplicative data augmentation

The methodology augments along three procedural axes:

- **Spatial conjugation**: Automatically synthesizing thousands of new conjugations from infinitives by mapping them to target spatial loci, with an inverse kinematics algorithm [3].
- **Prosodic variation**: Modifying temporal dynamics (tempo, holds, and easing) within linguistically attested ranges [4].
- **Body morphology**: Scaling body proportions (arm length, shoulder width) to simulate inter-signer diversity [5] and prevent model overfitting.

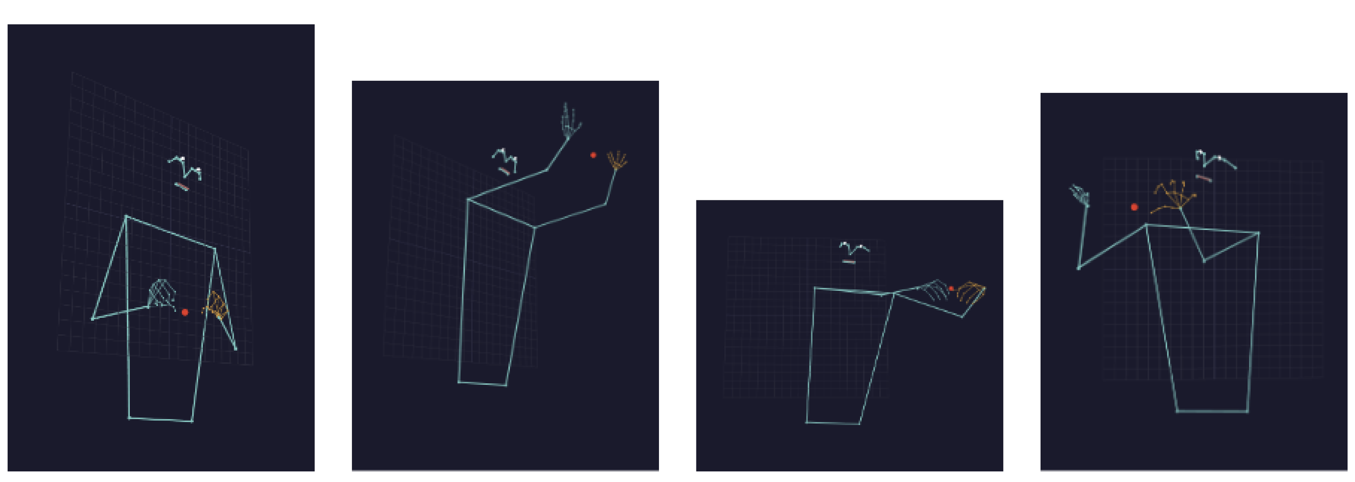
Augmented dataset size as a function of *V* (number of infinitive verbs) and *n* (number of prosodic x morphological variants:

V	Base ($\times$13)	n=5	n=10	n=25	n=50
1	13	65	130	325	650
10	130	650	1,300	3,250	6,500
25	325	1,625	3,250	8,125	16,250
48	624	3,120	6,240	15,600	31,200

5. Evaluation and conclusions

- **Three-layered validation**: The system is assessed based on anatomical feasibility (not exceeding physiological limits), kinematic fluency (smooth velocity curves), and linguistic legibility (user testing with deaf LSC signers).
- **Key finding**: Preliminary results suggest that procedural synthesis can effectively bridge the data gap for low-resource sign languages.
- **Future work**: Integration of these synthetic datasets into end-to-end SLR systems to measure actual improvements in translation performance.

Examples of procedurally generated skeletons using different spatial conjugations from the same infinitive verb TELL:



References

[1] Bragg, Danielle et al. 2019. Sign language recognition, generation, and translation: An interdisciplinary perspective. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*.

[2] Barberà, Gemma. 2015. *The meaning of space in sign language. Reference, specificity and structure in Catalan Sign Language discourse*. Mouton de Gruyter & Ishara Press, Berlin/Boston.

[3] Aristidou, Andreas et al. 2018. A. Inverse kinematics techniques in computer graphics: A survey. Computer Graphics Forum.

[4] Wilbur, Ronnie. 2009. Effects of varying rate of signing on asl manual signs and nonmanual markers. Language and speech, 52

[5] Li, Dongxu et al. 2020. Transferring cross-domain knowledge for video sign language recognition. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)

Funded by the European Union Next Generation EU and supported by the Government of Catalonia Grant Agreement No. SDC007/25/000093.

Finançat per la Unió Europea NextGenerationEU

Gobierno de España

Pla de Recuperació, Transformació i Resiliència

Generalitat de Catalunya Departament d'Empresa i Treball