Towards Integrating Pose Estimation with Neuroimaging for the Analysis of Signed Language Video Stimuli

Sébastien Vandenitte[1], Doris Hernández[2], Jarkko Keränen[2], Tommi Jantunen[2], Anna Puupponen[2]
[1]University of Namur, LSFB-Lab; [2]University of Jyväskylä, Sign Language Centre
[1]Rue de Bruxelles 61, B-5000 Namur, Belgium, [2]P.O. Box 35, FI-40014 University of Jyväskylä, Finland
sebastien.vandenitte@unamur.be, {doris.m.hernandez-barros, jarkko.j.keranen, tommi.j.jantunen, anna.m.puupponen}@jyu.fi

Abstract

We present our project revisiting the video stimuli of an EEG study in Finnish Sign Language to ask whether kinematic properties of the videos impacted their processing by study participants. For each stimulus, an average measure of brain responses across participants is computed. To analyse movement properties in the video stimuli, we rely on MediaPipe for pose estimation. We subsequently report on our project to perform an exploratory analysis of the kinematic properties of the videos which may affect their processing. We focus on several landmarks: the signer's right and left wrists, nose, and upper torso. Our goal is to obtain a kinematic profile of each stimulus video using several average kinematic variables: velocity and acceleration for all selected landmarks, distance between the wrists, and surface covered by the triangular area defined by the left hand, the right hand, and the nose. We conclude by discussing the potential benefits and limitations of this methodological approach.

Keywords: pose estimation, EEG, kinematics

1. Introduction

Quantifying and analyzing motion characteristics in signing is crucial in signed language research, particularly in approaches that emphasize the importance of strong connections between linguistic theory and phonetic, psychological, or neurological knowledge (Bybee, 1999; Diessel, 2017). The necessity of conducting kinematic analyses also extends to materials used for studying the neural processing of signed languages. Using video stimuli in neuroimaging studies of signed language processing is crucial for achieving a naturalistic and ecologically valid approach (see, e.g., Hosemann et al., 2013; Hernandez et al., 2022b; 2024; 2025). However, this creates a trade-off, as videos also introduce challenges for brain analysis. Compared to still images, signed language video stimuli are inherently complex and create a less controllable experimental environment. As a result, most features of the video signal have traditionally been treated as noise and excluded from analysis. Yet doing so results in the loss of important information about the dynamics of a visual-kinesic language. Addressing these typically ignored features requires controlling human motion in ways that produce systematically distinct conditions—an inherently difficult task, given the natural variability of human movement (Hernandez et al., 2022b).

Quantifying human motion in video stimuli through motion-capture or pose-estimation methods may offer solutions for improving control over the conditions studied. Such approaches can support the analysis of neural variables and aid in developing methodologies for using video materials in future neuroimaging studies on signed languages (as well as in non-neuro experimental settings). One way to achieve this is by applying openly available pose-estimation tools to characterize the kinematic properties of signed video materials. The pose data can then be used to examine the interaction between the kinematic characteristics of the signed stimuli and the brain's processing of those stimuli.

To the best of our knowledge, there are few studies that combine neuro-based analyses of signed language processing with kinematic analysis of signed language materials used as stimuli. Rivolta et al. (2025) first recorded short signed narratives with a Kinect set-up, thereby acquiring videos as well as motion-capture measurements of the stories. They used this information to extract time series of velocity for five articulators, namely the right hand, the left hand, the head, the torso, and a combined variable of the aforementioned articulators ("whole signal"). Next, they used magnetoencephalography (MEG) recordings to analyze neuronal oscillations of participants exposed to the signed narratives as stimuli. Rivolta et al. then converted the MEG data to the sampling rate of the motion-tracking data, enabling an analysis of how MEG data and articulator speed were coupled.

In this paper, we describe ongoing work in which we use MediaPipe pose estimation to analyse signed stimuli after they were used to acquire brain imaging data for a processing study. We describe how pre-existing brain-imaging data can be revisited to determine the potential effects of kinematic variables into the brain responses already obtained. Rather than synchronizing the brain data and the pose data, our approach relies on the association between kinematic variable

Proceedings of the 12th Workshop on the Representation and Processing of Sign Languages, pages 477–483

averages, e.g. mean head velocity and average brain responses, e.g., mean amplitude of P3, for each stimulus video. We also discuss the benefits and limitations of this methodological integration beyond our research questions.

2. Methodological approach

The data and methods presented here build on our previous work on the processing of Finnish Sign Language (FinSL). This earlier work comprises video stimuli consisting of isolated 'signs', response-time indices, and EEG measurements collected in a detection task investigating signers' recognition of enactment in signs with varying degrees of this depictive strategy, as well as their behavioural responses to these stimuli (Hernandez et al., 2025). In the present paper, we describe how we conduct kinematic analysis on the existing FinSL video stimuli and set out to combine these analyses with the existing EEG measurements in order to investigate potential interactions between signers' processing of FinSL depiction and the kinematic features involved in that depiction.

2.1 Pose estimation

Pose-estimation technology has been a major methodological breakthrough for the study of multimodal communication, gesture research, and signed language linguistics. In addition to paving the way for new research questions, this technology also enables researchers to revisit previous research. Pouw et al. (2021) revisit the gestures filmed and manually annotated for a previous study by Motamedi et al. (2019) to produce a fine-grained analysis of gestures with pose estimation. Trettenbrein and Zaccarella (2021) point out that pose estimation provides a new way to control for kinematic differences in stimuli used to study the perception of gesture and signed languages:

> *OpenPoseR provides quantitative measures of motion based on velocity and acceleration of the actor in the video which can be used for controlling differences in these movement parameters, for example, between different conditions of an experiment. More precisely, the package makes it possible to straightforwardly compute the Euclidean norms of sums of all velocity or acceleration vectors* […] *and thereby provide a quantitative measure of motion for an entire video clip.*
>
> Trettenbrein and Zaccarella (2021, p. 2)

Properties of the visual signal have notably been used in studies on the processing of signed language articulation. Brookshire et al. (2017) quantify changes in the visual stimulus over time with the IVC (Instantaneous Visual Change), a "time-series of aggregated visual change between frames [...] computed as the sum of squared differences in each pixel across sequential frames of video" (2017, p. 6356). More recently, Rivolta et al. (2025) study the coherence between MEG data and motion-tracking information in signed videos, in particular the speed of the right- and left-hand, of the head, and of the whole upper body.

Similarly, we use pose estimation to revisit the video stimuli used in a neurocognitive study for a post-hoc analysis. This methodological approach is motivated by Hernández et al.'s (2025) study on the perception of Finnish Sign Language, which led the authors to posit that visual saliency and an increase in the number of enacting bodily articulators affect attentional resources allocated for the detection of enactment in the signed language signal. The video data comprised 120 video clips of signed expressions in FinSL. Figure 1 illustrates a stimulus set, each image still retrieved from a separate video. All four videos in a set refer to the same action (here, "typing on a keyboard") but they vary in their degree of enactment, the independent variable in Hernández et al. (2025).

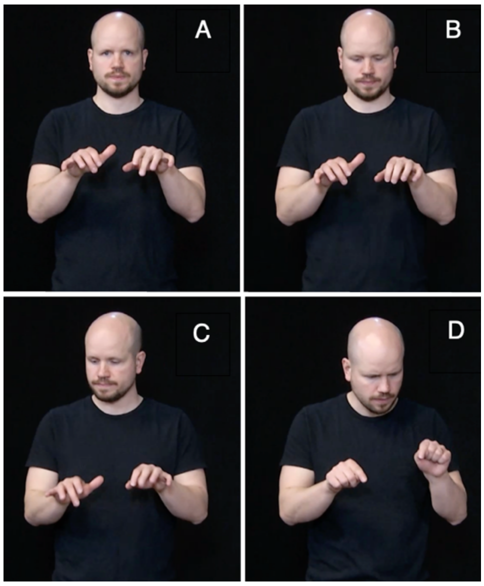

Figure 1: Stimulus set of 4 video stimuli with an increasing degree of enactment: no enactment (A), some enactment (B), substantial enactment (C), and full-blown enactment (D).

To analyse the kinematics of the stimuli, all 120 videos shown to participants were processed for pose estimation using MediaPipe Pose Landmarker (Lugaresi et al., 2019). Figure 2 illustrates the processing with original stills from a video stimulus (A) and stills from the same video processed for pose estimation (B). The dataframe generated after pose estimation contains three-dimensional coordinates for all landmarks. Given the recorded data is two-dimensional, the x and y coordinates are used whereas the z coordinates

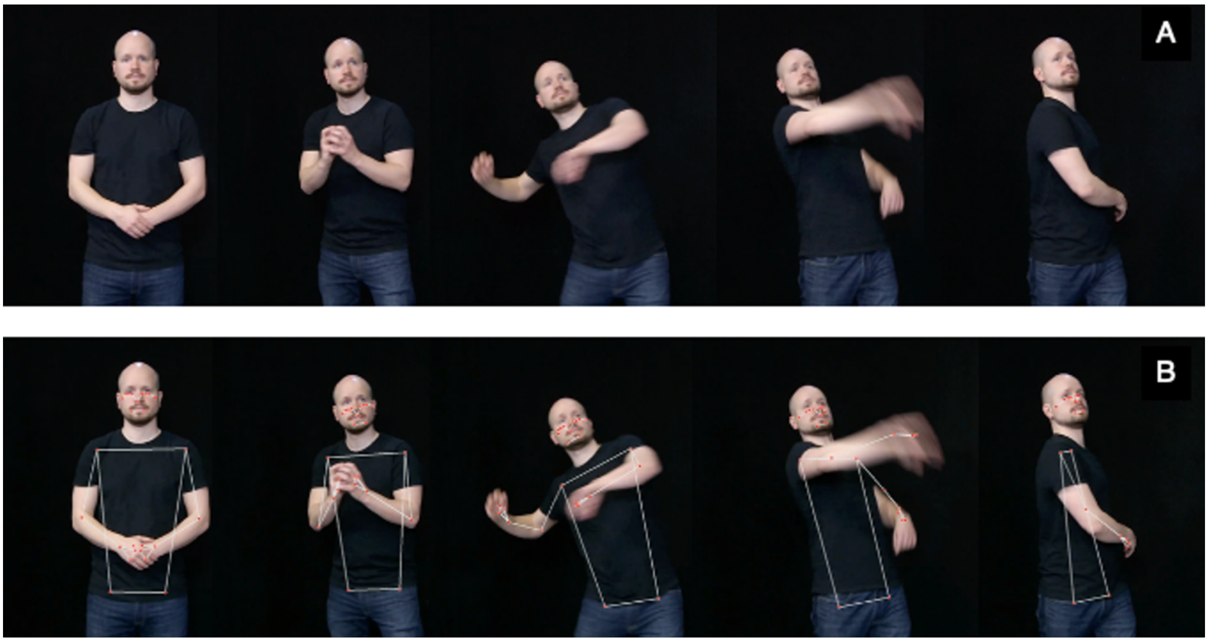

Figure 2: Original video stimulus (A) and video stimulus processed for pose estimation (B).

are discarded from the analysis, for reliability purposes (Ivashechkin et al., 2023). Next, landmarks corresponding to our articulators of interest are selected, namely the left and right wrist, and the nose. In addition, an upper-torso centroid landmark is generated as the mid-point between the left shoulder and the right shoulder. Our recording conditions provided good material for pose estimation: a single person is facing the camera in front of a dark background and stands at the same distance from the camera throughout the stimuli. The data processing proceeds with the generation of several kinematic variables based on the x and y coordinates.[1] These variables will be selected for their attested relevance in signed language phonetics while keeping a wide scope for the purpose of an exploratory approach:

- velocity (nose, upper torso, left wrist, right wrist)
- acceleration (nose, upper torso, left wrist, right wrist)
- distance between the left and right wrists
- surface of the area delimited by the nose, the left wrist, and the right wrist

The selection of these variables is motivated by previous studies. First, Jantunen et al. (2020) have shown significant differences in the velocity and acceleration of head and torso movements in motion-capture analyses of the phenomenon that we focus on. It remains to be seen whether these differences also affect processing. Second, we follow Rivolta et al. (2025) in analysing both the motion of individual articulators (head, left hand, right hand, torso) and the combined motion of varied articulators (here, the surface of the area within the perimeter surrounding the nose, the left wrist and the right wrist). Together, these variables provide us with complementary indices of the saliency of the motion signal, by measuring time derivatives and the size of the signing space. It is this saliency that Hernández et al. (2025) posited to affect participants' detection of enactment in the same video stimuli.

2.2 Processing of brain-based functional data

In neuroscience, several techniques can be used to examine either the structural (anatomical) or functional (operational) properties of the brain. Structural techniques, such as Magnetic Resonance Imaging (MRI), and functional techniques, such as Electroencephalography (EEG) and Magnetoencephalography (MEG), can both benefit from being combined with pose data. In this paper, we focus on EEG measurements of electrical activity, reflected as voltage changes arising from the synchronized post-synaptic potentials of large groups of neurons.

During the acquisition of EEG data, electrical activity is recorded from multiple electrodes placed on standardized scalp locations covering all the scalp. Following preprocessing (see Puce and Hämäläinen, 2017), the continuous EEG signal is divided into segments, or epochs, surrounding the events of interest (e.g. the onset of a signed language video). Each epoch usually includes a pre-stimulus baseline period to allow for the correction of voltage differences before the stimulus is presented. By aligning these epochs across multiple trials, the EEG signals for each stimulus can be averaged.

[1] To tackle errors in landmark detection and reduce noise in the generated kinematic variables, we apply Savitzky–Golay smoothing filters to the obtained coordinates as well as in the computation of time derivatives (Pouw, 2024).

File name	Mean nose velocity	Mean RH acceleration	…	Mean LH-RH distance	Mean surfacre: LH-RH-nose area	Mean N2 amplitude	Mean P3 amplitude
video 1	0.0106	3.1114	…	0.1492	0.0090	-0,4250	2,5329
video 2	0.0252	4.2351	…	0.2369	0.0112	-1,1591	3,0462
video 3	0.0064	3.1789	…	0.2673	0.0087	-1,4625	1,0295
video 4	0.0367	5.1376	…	0.1907	0.0042	-0,8466	0,5127

Table 1: Illustration of the data matrix combining the pose-estimation and EEG analysis.

This process enhances activity that is consistently synchronised with the stimulus, while reducing unrelated background activity. Instead of taking all epochs per person and averaging them, in the approach documented here we average the signal acquired for all subjects per epoch, in order to triangulate it with the pose-estimation analysis. In our experiment, as each stimulus was a video containing a single sign, the epochs were time-locked to the onset of the video. Therefore, an epoch represents a segment of the EEG signal that reflects the brain's response to a specific sign. Each epoch included a 200 ms pre-stimulus baseline period prior to video onset, followed by an 800 ms post-stimulus time window.

Epochs can be further used in several ways. One of these ways is to analyze the EEG data in the time domain by using event-related potentials (ERP). ERP is the term given to voltage fluctuations in the brain's electrical activity that occur in response to a particular sensory, cognitive or motor stimulus (like the presentation of a signed language video). ERPs are particularly useful as they enable researchers to examine the timing and sequence of neural events with millisecond precision (for a review of ERPs in research on signed languages, see Hernández et al., 2022a). Another way of further using epochs is by analyzing EEG data in the frequency domain by decomposing the signal into brain oscillations – rhythmic patterns in electrical brain activity. For one or several oscillations (i.e. theta, alpha, beta, gamma, etc.), several properties are commonly extracted (e.g., power, phase, bandwidth). They reflect the brain's response to the experimental conditions (for a review, see Drijvers and Mazzini, 2023).

Traditionally, EEG analyses have focused on temporal/frequency (ERPs, oscillations) patterns across trials. Individual epochs are often merely treated as inputs to these averages, with relatively little attention given to variability across trials or within single epochs (example Hernández et al., 2024; 2025). Instead of this, in the approach described here we focus on the information contained within each individual epoch rather than relying exclusively on their aggregated averages. In this approach, data from multiple subjects are combined at the group level, averaging across participants while maintaining a detailed representation of the neural responses in each epoch. This enables a more nuanced exploration of the neural processes associated with each stimulus presentation and captures subtle temporal (latency) or size (amplitude/power) variations that are usually overlooked in conventional EEG analyses.

2.3 Binding pose estimation and brain functional data

After pre-processing the pose-estimation and the EEG data, the database created for the analysis includes columns representing each extracted measure of kinematic and brain-derived variables (Table 1). Each row corresponds to one video presenting a signed language stimulus, rather than a participant, as is usually the case in EEG studies. In this dataframe, a column for each kinematic variable provides average measurements for each stimulus video, e.g., average nose velocity or average distance between the left and the right wrist. In addition, the database includes columns for the EEG variables, which provide average measurements of the neural responses to each stimulus across participants (per epoch), as described in Section 2.2. The specific variables in Table 1 (N2 and P3 ERPs) are motivated by our previous work and based on the processing and analysis of these components in that study (Hernández et al., 2025).

Once the potential associations have been identified, we plan to use linear mixed models (LMMs) to confirm their statistical robustness, because of their flexibility to include fixed and random effects and to control for covariates (West et al., 2022). The idea behind the confirmatory analysis is to know how kinematic variables explain part of the variance in the brain variables. Figure 3 demonstrates a typical workflow for the methodological integration presented here, modelled after our goal to study the processing of FinSL videos.

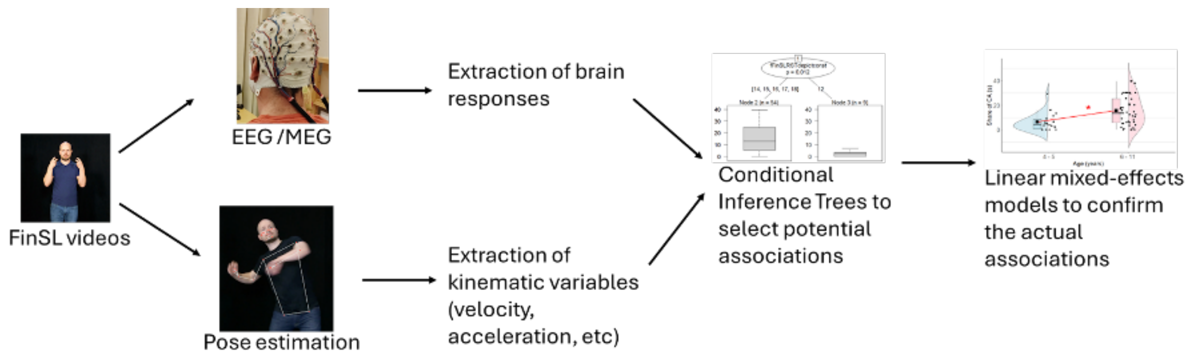

Figure 3: Diagram showing the study design and methodological approach suggested.

3. Discussion

In this paper, we have presented the combination of two methodological tools often used in different research traditions: motion tracking and brain imaging. Using fine-grained information from pose estimation enables joining brain analyses to visual information about the stimuli with an unprecedented level of detail. In what follows, we suggest some research avenues that could benefit from this methodological approach.

3.1 Benefits and application of the approach

Using kinematic analysis of signed language video stimuli is a methodological development that is beneficial for cognitive neuroscience. Pose-estimation methods can be used to characterize the videos (to be) used in neurocognitive studies. When used for post-hoc analysis, they provide detailed information about the movement properties of the stimuli. A direct benefit of this information is that it can be exploited to check for the effects of motion properties on the processing of the stimuli. In the long run, the variability in movement across stimuli could also be taken into account like other visible factors in video materials, such as background or luminance have been taken into consideration in previous studies (e.g. Hernandez et al., 2022b). This, in turn, can facilitate the building of video stimuli for future neurolinguistic studies on the processing of signed languages, enabling higher degrees of stimulus control and isolation of the target phenomena.

In addition, the increased explanatory power derived from pose estimation also facilitates pursuing several research questions that have emerged in signed language phonetics. It enables asking whether we find neural correlates to some of the phonetic phenomena studied or posited by signed language linguists. For instance, do differences in velocity and acceleration between signs and sign transitions lead to processing differences, notably at the level of perception (Jantunen, 2013)? Do we find neural evidence of a different processing of highly lexical signs as against depicting signs given their different kinematic signatures (Stamp et al., 2024)? If so, what is the nature of these differences?

Finally, another benefit of this methodological combination is its sustainability. Indeed, granted open access to such datasets, it enables conducting further research without having to use new resources for data collection. The maximal use of pre-existing materials is supported by current data protection policies (e.g., General Data Protection Regulation) as well as international guidelines for data management such as the FAIR principles (Wilkinson et al. 2016).

3.2 Limitations of the approach

Pose estimation and other forms of video-based motion-tracking offer clear advantages over methods that require specialized hardware, cameras and studio/laboratory environments, such as optoelectronic motion capture. These advantages include greater accessibility, lower cost, and the possibility of analysing pre-existing signed language video materials. However, these benefits come with a trade-off in accuracy, as such approaches do not achieve the same level of detail or precision as motion-capture methodologies in analyzing signed language kinematics (Bauer et al., 2024; Bux et al., 2024). This limitation affects the reliability of the resulting kinematic measurements when compared with those obtained through more specialized techniques. Nevertheless, the accuracy of pose estimation is expected to improve as the underlying technologies continue to develop.

In addition, the methodology presented here relies on average (kinematic and neural) measures associated with each stimulus, without synchronizing the EEG data to the estimated kinematic signals. In future work, the synchronization of the EEG data to the estimated kinematics signals is an important task (Gregori et al., 2023), though not a simple one. While EEG operates at a millisecond level, pose estimation is based on the framerate of videos. This can cause problems with latency measures, especially when working with ERPs or time-frequency representations (TFRs) that need to be strictly locked to stimulus timing. At the moment, this can be done, for example, with Power Spectral Density (PSD) analysis (Rivolta et al., 2025) which looks at how the power of a signal is distributed across different frequencies disregarding time.

Finally, the exploratory study proposed in this paper is limited in scope. On the one hand, the dimensions of signed language motion suggested for analysis in this paper only capture part of the signed language motion signal. Other articulators in the stimuli, e.g., eye gaze or mouth actions, could also play a role in the processing of the videos. In the same vein, other kinematic features may also be of relevance. On the other hand, this study focuses on the potential influence of kinematic features on the neural processing of signed language video stimuli. However, further development of this approach should consider additional differences between the videos that may impact their processing, such as differences in sign meaning or degree of perceived iconicity.

4. Conclusion

In this paper, we have presented our experience combining EEG data with the pose data of the video stimuli to which study participants were exposed. After presenting pose estimation and the use of EEG for brain imaging, we exemplified one way of combining these datasets by using stimulus/epoch-specific averages of kinematic and neural activity variables. While pose estimation exhibits some limitations in terms of accuracy and relative schematicity, it provides a low-resource method to obtain kinematic information about the video stimuli. This methodological combination enables us to perform a post-hoc analysis of how motion properties of the video stimuli affected their processing, as measured by EEG.

5. Acknowledgments

This research was supported by the Research Council of Finland (ShowTell project, Grant No. 339268).

DH acknowledges funding from the Finnish Cultural Foundation - SKR (Grant No. 00250359).

SV is a Postdoctoral Researcher of the *Fonds de la Recherche Scientifique – FNRS*.

6. Bibliographical References

Bauer, A., Kuder, A., Schulder, M. and Schepens, J. (2024). Phonetic differences between affirmative and feedback head nods in German Sign Language (DGS): A pose estimation study. *Plos one,* 19(5), e0304040.

Brookshire, G., Lu, J., Nusbaum, H. C., Goldin-Meadow, S. and Casasanto, D. (2017). Visual cortex entrains to sign language. *Proceedings of the National Academy of Sciences,* 114(24), 6352-6357.

Bux Sargano, A., Vandenitte, S., Jantunen, T. and Kimmelman, V. (2024). Evaluation of Head Pose Estimation Algorithms for Sign Language Analysis. In *International Conference on IT and Industrial Technologies (ICIT)*, pages 1-6, Chiniot, Pakistan. https://doi.org/10.1109/ICIT63607.2024.10859486

Bybee, J. (1999). Usage-based phonology. *Functionalism and formalism in linguistics*, 1, 211-242.

Diessel, H. (2017). Usage-based linguistics. In Mark Aronoff (Ed.), *Oxford Research Encyclopedia of Linguistics*. Oxford University Press, Oxford, United Kingdom. https://doi.org/10.1093/acrefore/9780199384655.013.363

Drijvers, L., and Mazzini, S. (2023). Neural oscillations in audiovisual language and communication. In *Oxford Research Encyclopedia of Neuroscience*. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780190264086.013.455

Gregori, A. et al. (2023). A Roadmap for Technological Innovation in Multimodal Communication Research. In: Duffy, V.G. (eds) *Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management.* HCII 2023. Lecture Notes in Computer Science, vol 14029, pages 402-438. Springer, Cham. https://doi.org/10.1007/978-3-031-35748-0_30

Hernández, D., Puupponen, A., Keränen, J., Vandenitte, S., Anible, B., Ortega, G. and Jantunen, T. (2025). Neuroelectrical and behavioral correlates of constructed action recognition in Finnish sign language. *Neuroscience,* 575, 140-149. https://doi.org/10.1016/j.neuroscience.2025.03.046

Hernández, D., Puupponen, A., Keränen, J., Ortega, G., and Jantunen, T. (2024). Between bodily action and conventionalized structure : The neural mechanisms of constructed action in sign language comprehension. *Brain and Language,* 252, Article 105413. https://doi.org/10.1016/j.bandl.2024.105413

Hernández, D., Puupponen, A., and Jantunen, T. (2022a). The contribution of event-related potentials to the understanding of sign language processing and production in the brain: Experimental evidence and future directions. *Frontiers in Communication,* 7, 750256. https://doi.org/10.3389/fcomm.2022.750256

Hernández, D., Puupponen, A., Keränen, J., Wainio, T., Pippuri, O., Ortega, G. and Jantunen, T. (2022b). Use of Sign Language

Videos in EEG and MEG Studies : Experiences from a Multidisciplinary Project Combining Linguistics and Cognitive Neuroscience. In J. H. Jantunen, J. Kalja-Voima, M. Laukkarinen, A. Puupponen, M. Salonen, T. Saresma, J. Tarvainen, and S. Ylönen (Eds.), *Diversity of Methods and Materials in Digital Human Sciences: Proceedings of the Digital Research Data and Human Sciences DRDHum Conference 2022*, December 1-3, pages 148-155. Jyväskylä, Finland. Jyväskylän yliopisto. http://urn.fi/URN:ISBN:978-951-39-9450-1

Hosemann, J., Herrmann, A., Steinbach, M., Bornkessel-Schlesewsky, I. and Schlesewsky, M. (2013). Lexical prediction via forward models: N400 evidence from German Sign Language. *Neuropsychologia,* 51(11), 2224-2237. https://doi.org/10.1016/j.neuropsychologia.2013.07.013

Hothorn, T., Hornik, K. and Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical statistics,* 15(3), 651-674.

Ivashechkin, M., Mendez, O. and Bowden, R. Improving 3D Pose Estimation For Sign Language. In: *2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW);* 2023, pages 1–5. https://doi.org/10.1109/ICASSPW59220.2023.10193629

Jantunen, T. (2013). Signs and transitions: Do they differ phonetically and does it matter?. *Sign Language Studies,* 13(2), 211-237.

Jantunen, T., Weerdt, D., Burger, B. and Puupponen, A. (2020). The more you move, the more action you construct: A motion capture study on head and upper-torso movements in constructed action in Finnish Sign Language narratives. *Gesture,* 19(1), 72–96. https://doi.org/10.1075/gest.19042.jan

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C., Yong, M.G., Lee, J., Chang, W., Hua, W., Georg, M. and Grundmann, M. (2019). MediaPipe: A Framework for Building Perception Pipelines. *ArXiv, abs*/1906.08172.

Motamedi, Y., Schouwstra, M., Smith, K., Culbertson, J. and Kirby, S. (2019). Evolving artificial sign languages in the lab: From improvised gesture to systematic sign. *Cognition,* 192, 103964.

Pouw, W., Dingemanse, M., Motamedi, Y. and Özyürek, A. (2021). A Systematic Investigation of Gesture Kinematics in Evolving Manual Languages in the Lab. *Cognitive Science,* 45(7), e13014. https://doi.org/10.1111/cogs.13014

Pouw, W. (2024). Wim Pouw's EnvisionBOX modules for social signal processing (Version 1.0.0) [Computer software]. https://github.com/WimPouw/envisionBOX_modulesWP

Puce, A. and Hämäläinen, M. S. (2017). A review of issues related to data acquisition and analysis in EEG/MEG studies. *Brain sciences,* 7(6), 58.

Rivolta, C. L., Lizarazu, M., Costello, B., and Carreiras, M. (2025). Cortical tracking of sign language: The role of language knowledge in tracking of different articulators. *Proceedings of the National Academy of Sciences*, 122(51), e2512665122.

Stamp, R., Cohn, D., Hel-Or, H., and Sandler, W. (2024). Kinect-ing the Dots: Using Motion-Capture Technology to Distinguish Sign Language Linguistic From Gestural Expressions. *Language and Speech,* 67(1), 255–276. https://doi.org/10.1177/00238309231169502

Trettenbrein, P. C., and Zaccarella, E. (2021). Controlling video stimuli in sign language and gesture research: The OpenPoseR package for analyzing OpenPose motion-tracking data in R. *Frontiers in Psychology,* 12, 628728.

West, B. T., Welch, K. B., and Galecki, A. T. (2022). Linear mixed models: a practical guide using statistical software. Chapman and Hall/CRC. https://doi.org/10.1201/9781003181064

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 https://doi.org/10.1038/sdata.2016.18