

Responsible NLP Checklist

Paper title: *Selective Contrastive Learning For Gloss Free Sign Language Translation*

Authors: *Chang hao lai, Rui Zhao, Xuwen Zhong, Jinsong Su, Yidong Chen*

How to read the checklist symbols:

- the authors responded 'yes'
- the authors responded 'no'
- the authors indicated that the question does not apply to their work
- the authors did not respond to the checkbox question

For background on the checklist and guidance provided to the authors, see the [Responsible NLP Checklist](#) page at ACL Rolling Review.

A. Questions mandatory for all submissions.

A1. Did you describe the limitations of your work?
This paper has a Limitations section.

A2. Did you discuss any potential risks of your work?
Ethics Statement

B. Did you use or create scientific artifacts? (e.g. code, datasets, models)

B4. Did you discuss the steps taken to check whether the data that was collected/used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect/anonymize it?
We utilized standard, publicly available benchmark datasets (PHOENIX14T and CSL-Daily) where participant consent and initial data vetting were handled by the original dataset creators. Furthermore, in Sign Language Translation (SLT), facial expressions and non-manual features are essential linguistic articulators. Visually anonymizing the subjects (e.g., via face blurring) is not technically feasible, as it would destroy the core linguistic signals required for this research.

B6. Did you report relevant statistics like the number of examples, details of train/test/dev splits, etc. for the data that you used/created?
Section 4.1

C. Did you run computational experiments?

C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?
Section 4.2

C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?
Due to the high computational cost and large memory footprint associated with spatio-temporal video modeling in Sign Language Translation, running multiple identical experiments with different random seeds to compute variance or error bars is computationally prohibitive. Therefore, following standard practice in recent video-based SLT research, we report the results based on a single representative run.

The [Responsible NLP Checklist](#) used at ACL Rolling Review is adopted from [NAACL 2022](#), with the addition of [ACL 2023](#) question on AI writing assistance and further refinements based on ARR practice. [ACL 2026](#) used a subset of ARR checklist form.

D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?

D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?

(left blank)

D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?

(left blank)

D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating (e.g., did your instructions explain how the data would be used)?

(left blank)

D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?

(left blank)

E. Did you use AI assistants (e.g., ChatGPT, Copilot) in your research, coding, or writing?

E1. If you used AI assistants, did you include information about their use?

We utilized AI assistants (e.g., ChatGPT) strictly for language polishing and grammar correction