

Marc Schulder<sup>1</sup>, Sam Bigeard<sup>1,2</sup>, Maria Kopf<sup>1</sup>, Thomas Hanke<sup>1</sup>, Anna Kuder<sup>3</sup>, Joanna Wójcicka<sup>4</sup>, Johanna Mesch<sup>5</sup>, Thomas Björkstrand<sup>5</sup>, Anna Vacalopoulou<sup>6</sup>, Kyriaki Vasilaki<sup>6</sup>, Theodore Goulas<sup>6</sup>, Stavroula–Evita Fotinea<sup>6</sup>, Eleni Efthimiou<sup>6</sup>

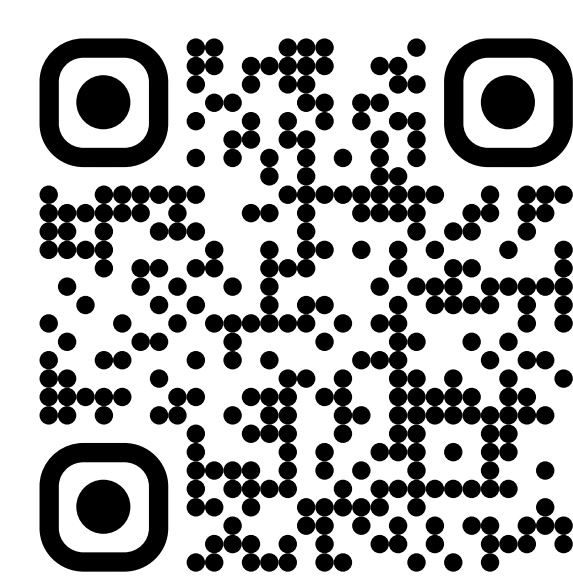
<sup>1</sup>Institute of German Sign Language and Communication of the Deaf, University of Hamburg, Germany | <sup>2</sup>Inria Centre, University of Lorraine, France  
<sup>3</sup>Department of Linguistics, University of Cologne, Germany | <sup>4</sup>Department of General Linguistics, Sign Language Linguistics and Baltic Studies, University of Warsaw, Poland  
<sup>5</sup>Department of Linguistics, Stockholm University, Sweden | <sup>6</sup>Institute for Language and Speech Processing, Athena Research Center, Greece

## The Multilingual Sign Language Wordnet

- Lexical resources with focus on meaning disambiguation
  - Annotated by language experts
- Custom annotation interface**
- Regularly updated based on annotator feedback
  - Automatic suggestions speed up annotation by indicating likely matches
- Builds on existing resources**
- Synsets (= meanings) from Open Multilingual Wordnet (OMW)
  - Sign inventory & videos from numerous lexical SL resources

### 8 sign languages

- ◆ Greek Sign Language (GSL)
- ◆ Polish Sign Language (PJM)
- ◆ Swedish Sign Language (STS)
- British Sign Language (BSL)
- German Sign Language (DGS)
- Swiss German Sign Language (DSGS)
- French Sign Language (LSF)
- Sign Language of the Netherlands (NGT)



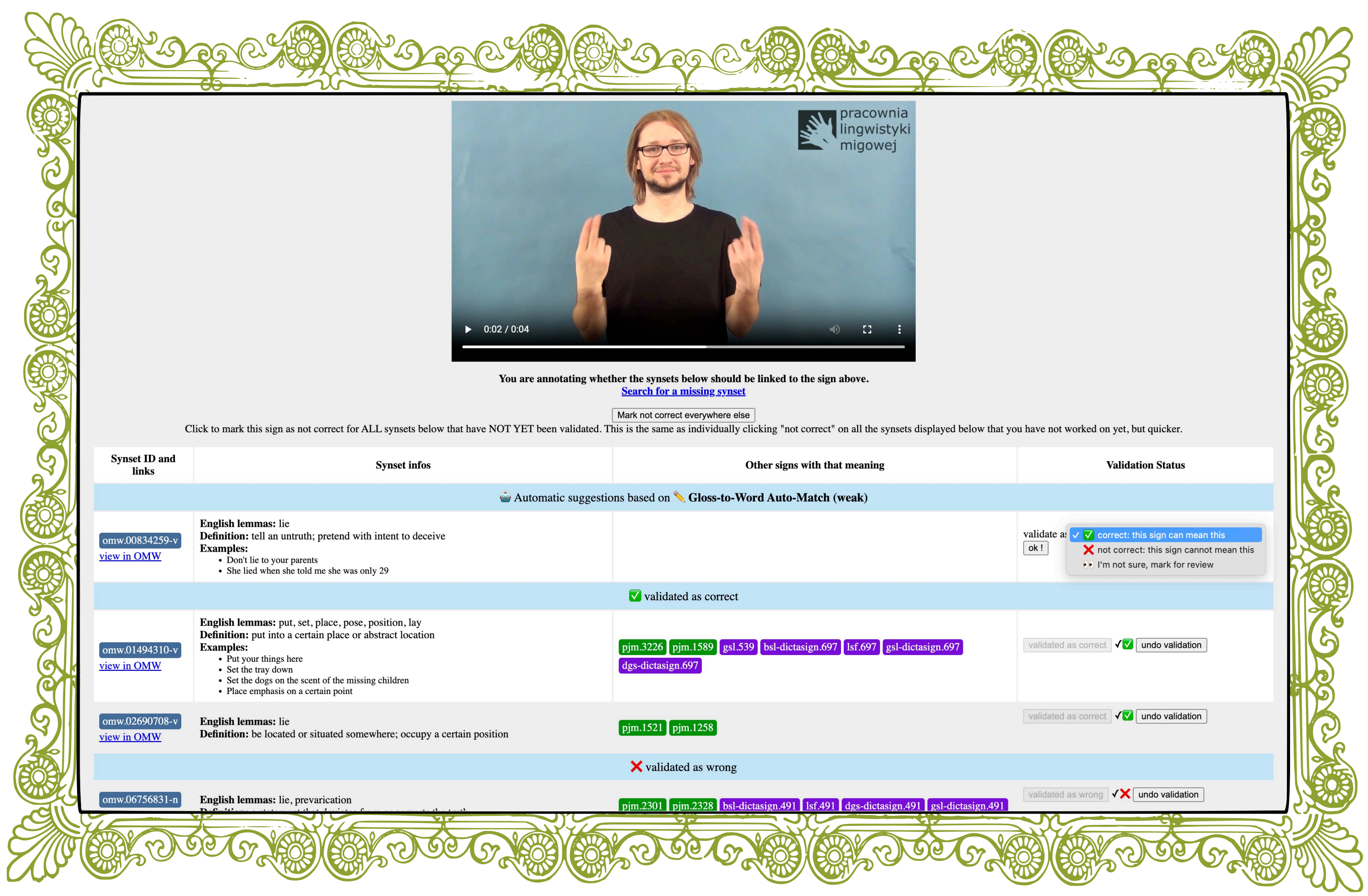
<https://sign-net.meine-dgs.de>

## Gloss-based suggestions

- Match sign glosses/keywords to OMW lemmas, copy synset links from lemmas
- **Advantages:** Easy to implement, required data is available from the start
- **Disadvantages:** *Form-based reverse translation*
  - Over-generates for polysemous words
  - Misses sign meanings not covered by gloss/keywords
  - Focus on spoken language information

## Annotation Interface

- Information**
- **Signs:** Video & gloss/keyword
  - **Synsets:** Definitions, examples & lemmas in English or regional spoken lang.
  - **Related annotations:** Show items that are already linked to an entry
- Perspectives**
- **Sign view:** Annotate all synsets of a single sign
  - **Synset view:** Annotate all signs of a single synsets



## New feature: Synonym-based suggestions

- Suggest meanings for signs, based on their synonymy with other signs:
  1. Look at completed sign-synset annotations
  2. Find signs that share some meanings already (= link to same synset)
  3. If one sign has more meanings, suggest those for the other sign.
- **Advantages:**
  - Lexical unit reverse translations more robust than form-based reversal
  - Focus on sign language information
  - Can be done across different languages
  - Speeds up annotation of entries that are (dialectal) variants of each other
  - Marking links as invalid produces strong information on partial synonymy

### Current annotation task



— manually verified  
- - - - automatic suggestion  
suggestion is ✓ valid / ✗ invalid

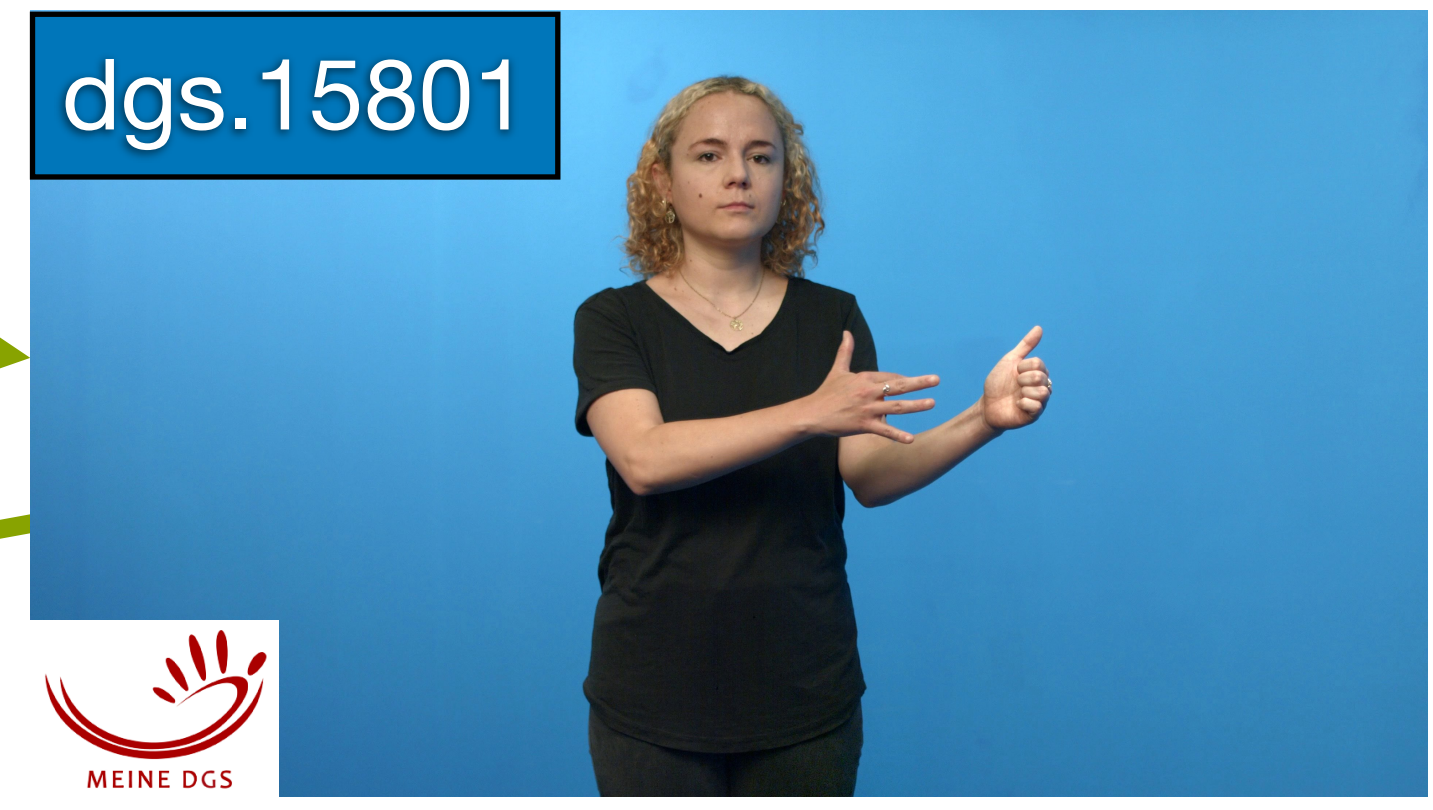
✓ omw.08859173-n  
Ireland (island)

✗ omw.00375969-a  
green

omw.08888676-n  
Republic of Ireland

✓ omw.03495258-n  
harp

### Completed annotations



## New Dataset Release

- New sign-synset links for GSL, PJM and STS
- Annotations assisted by new synonym-based suggestions
- Synsets covered by multiple sign languages:
  - 2+ sign languages: 3,361 synsets
  - 3+ sign languages: 1,472 synsets
  - 4+ sign languages: 617 synsets

	GSL	PJM	STS	All languages
Signs	2,949	2,415	2,706	14,028
Synsets	5,760	3,626	1,981	16,534
Links	7,499	6,486	2,810	24,367

Poster presented at the 11th Workshop on the Representation and Processing of Sign Languages (sign-lang@LREC-COLING 2024), Turin, Italy. 25 May 2024.

This publication has been produced in the context of the joint research funding of the German Federal Government and Federal States in the Academies' Programme, with funding from the Federal Ministry of Education and Research and the Free and Hanseatic City of Hamburg. The Academies' Programme is coordinated by the Union of the German Academies of Sciences and Humanities. This work is supported in part by the EASIER (Intelligent Automatic Sign Language Translation) Project. EASIER has received funding from the European Union's Horizon 2020 research and innovation programme, grant agreement n° 101016982. This work is supported in part by INRIA within the project Defi COLaF. AK is funded by the DFG Priority Programme 2392 Visual Communication (ViCom, 2022-2025), Frankfurt am Main, Germany.