

# Introducing Sign Languages to a Multilingual Wordnet: Bootstrapping Corpora and Lexical Resources of Greek Sign Language and German Sign Language

Sam Bigeard<sup>1</sup>, Marc Schulder<sup>1</sup>, Maria Kopf<sup>1</sup>, Thomas Hanke<sup>1</sup>, Kyriaki Vasilaki<sup>2</sup>, Anna Vacalopoulou<sup>2</sup>,  
Theodoros Goulas<sup>2</sup>, Athanasia-Lida Dimou<sup>2</sup>, Stavroula-Evita Fotinea<sup>2</sup>, Eleni Efthimiou<sup>2</sup>

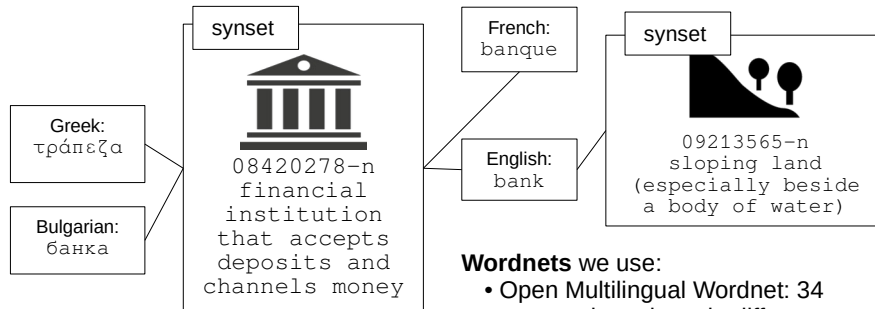
<sup>1</sup> University of Hamburg, Institute of German Sign Language and Communication of the Deaf, Germany

<sup>2</sup> Institute for Language and Speech Processing (ILSP), Athena Research Center, Greece

Beta version available at [sign-net.meine-dgs.de](http://sign-net.meine-dgs.de)



## What?



what we add



### Wordnets we use:

- Open Multilingual Wordnet: 34 connected wordnets in different languages
- Germanet: 28,564 synsets mapped to OMW
- Greek Wordnet: 18,049 synsets

### Sign resources we use:

- DGS Corpus
- Polytropon

## Why?

**Goal:** Connect existing sign language lexicons & make them machine-readable

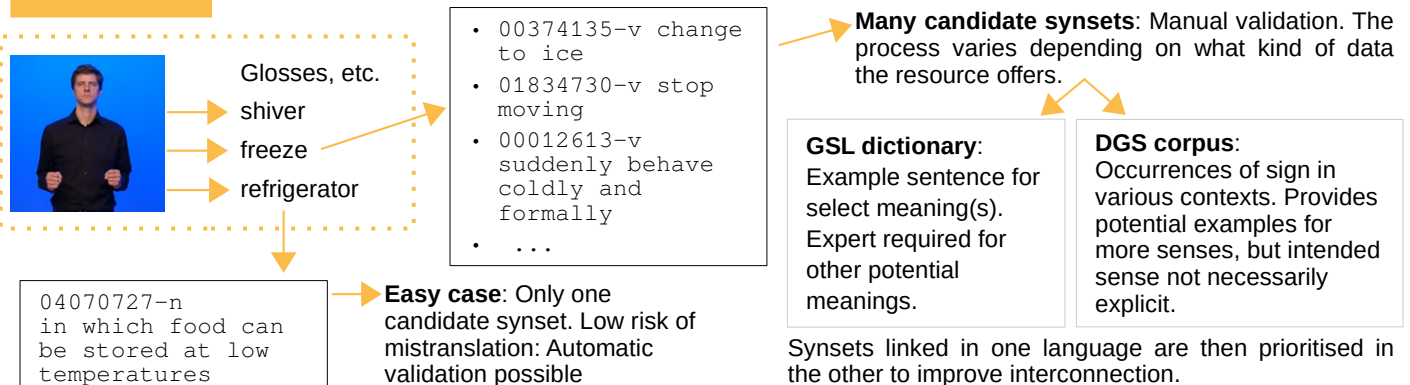
### Why a Wordnet?

- Precise senses, accessible through definition, words and relationships
- Sense is identified by number, not translation to spoken language
- Machine-readable. Libraries to manipulate it already exist

Why expand on existing wordnets rather than create one from scratch?

- Easier and quicker to build
- More accessible to future users
- Compatible with many resources: languages, pictures, etc.

## How?



## Current Progress

	GSL validated	DGS candidates	DGS validated	GSL/DGS overlap
distinct synsets	4214	27,020	969	278
distinct signs	1819	11,856	2230	N/A
sign-synset pairs	4347	138,518	2230	N/A

**validated:** ready to be used in applications  
**candidates:** automatic matches, needs to be verified  
**overlap:** synsets linked to both GSL and DGS

## Why are the ratios different?

**Corpus-based (DGS):** Many signs with overlapping meanings in corpus. One synset associated with a sign propagates to synonymous signs, resulting in more candidate signs.

**Dictionary-based (GSL):** Dictionary created with focus on providing many concepts with at least one sign. Results in faster validation of many synsets, but fewer signs per synset.

**Work in progress:** We expect the two datasets to evolve toward many sign-synset pairs, as is seen currently in the *DGS candidates* column.