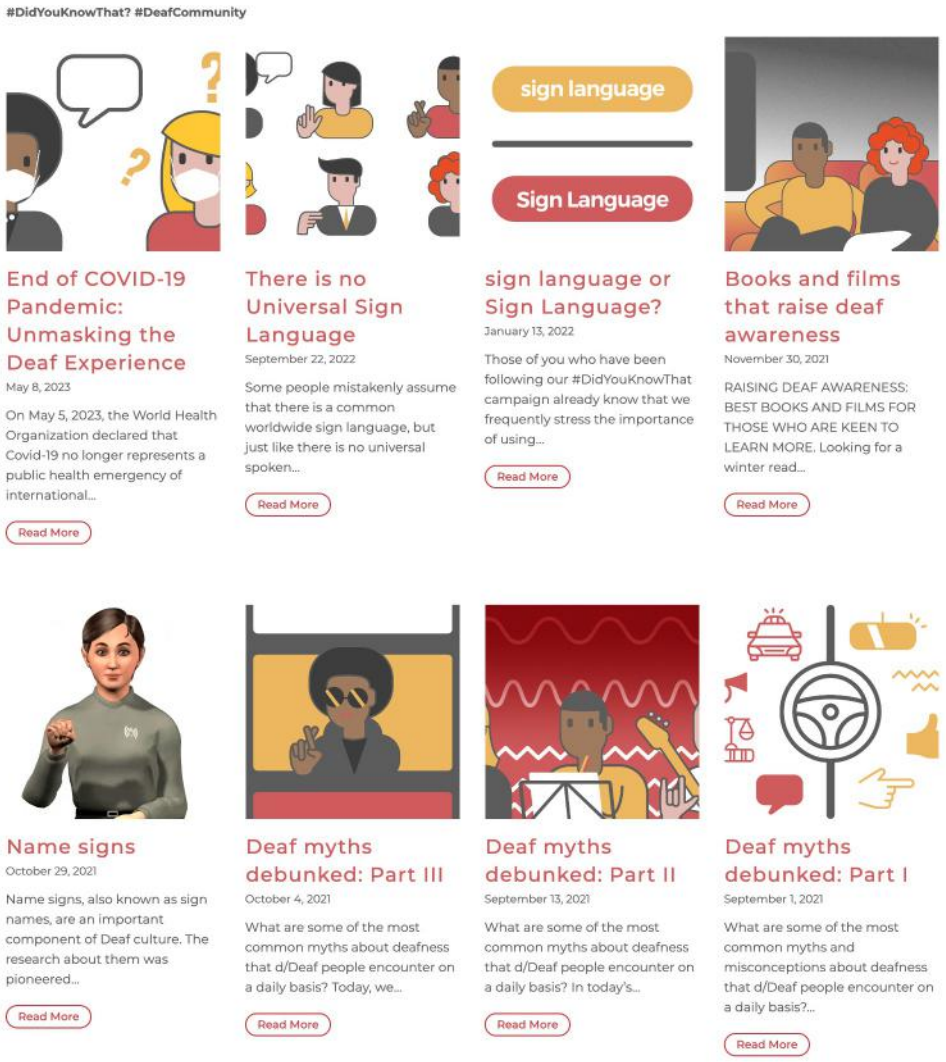
Sign Language and Sign Language processing technologies

Eleni Efthimiou

Research Director, ATHENA RC/ILSP

EASIER Scientific Coordinator

#DidYouKnowThat?



A few important things to know before starting:

1. There is no universal sign language!
2. Sign languages are living languages, just like spoken languages. Each Deaf community has their own language, e.g., BSL ≠ ASL ≠ English and DGS ≠ DSGS ≠ German.
3. Sign languages are not only about gestures. They inolve all body, facial expressions, and mouthing.
4. Spoken languages are foreign languages to native sign language signers.

Structure of the presentation

- Signed vs Spoken Language constructs in brief
- Technologies for processing SL
- Current state in technology and data development

Sign Language Constructs

Signed vs. Spoken languages

Spoken

Sentence

- Word
  - Phoneme

Signed

Utterance

- Sign
  - Phoneme

Both are defined through the identification of contrastive elements.

Single vs multimodal articulation

Spoken

Phoneme

One active parameter: **sound**

in linear sequences to form spoken language units

Signed

Phoneme

Five **simultaneously** active parameters to form signed language units:

- **handshapes,**
- **palm orientation,**
- **location,**
- **movement,**
- **nonmanual signals**

Modality

	Spoken	Signed
Production	Vocal apparatus	Concurrent movement of multiple body parts
Reception	Ear	Eye
Dictionary	Discrete lexical items	Forms with infinite gradations
Main structure	Linear sequence of lexical units	Simultaneity of forms Productive use of space
Written form	In many instances	None in widespread use

Context

- A sign in isolation has limited use (i.e. dictionary look-up)
- The context helps determine its meaning.
- Producing a signed sentence means putting signs in context
- Thus, modifying isolated sign forms

Context

- Dictates modifications to signs in citation form
- Result is an inflected sign that can differ in
  - Handshape
  - Palm Orientation
  - Location
  - Movement
  - Nonmanual signals

+ Timing + Prosody

=>

All 5 parameters are subject to change depending on CONTEXT

Indicative Inflection Examples

1. Posing questions
2. Negating
3. Size and shape specifiers
4. Proforms
5. Role Shift
6. Adverbial modifiers

In Summary

All linguistic parameters are subject to **multi-dimentional** change in context.

Timing is essential and is determined by context.

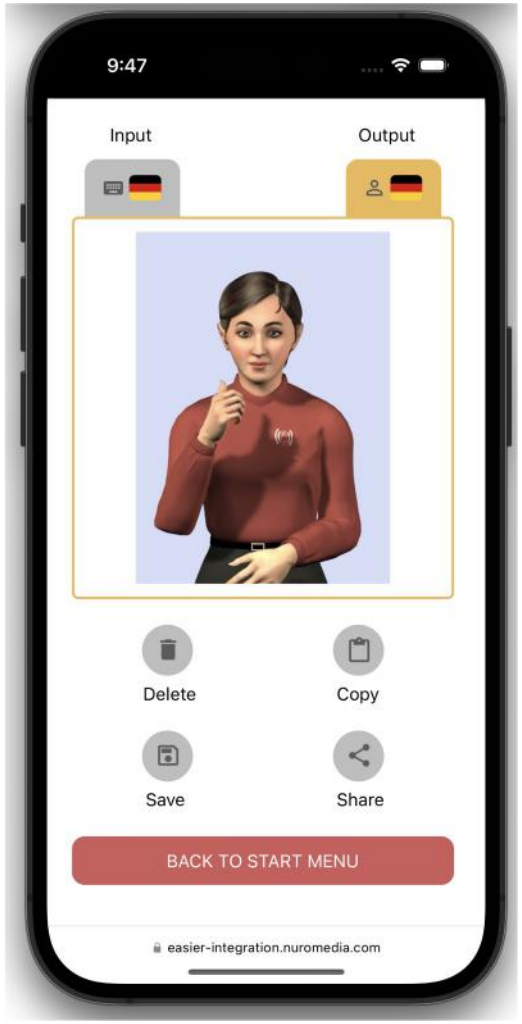
- *Within a sign*
- *Within an utterance*

Notation systems

	Machine readable	Inflection specification	Co-occurring processes	Data available
Annotated gloss stream	No	Some	Some	N/A
SignWriting	No	No	Some	N/A
Stokoe	No	No	No	N/A
HamNoSys	Yes	No	Some	Yes
Sign Language Phonetic Annotation	No	No	Some	Limited
Prosodic Model Handshape Coding	No	No	Some	Limited
SiGML	Yes	Some	Some	Yes
EMBRscript	Yes	No	Some	No
AZee	Yes	Yes	Yes	Not yet

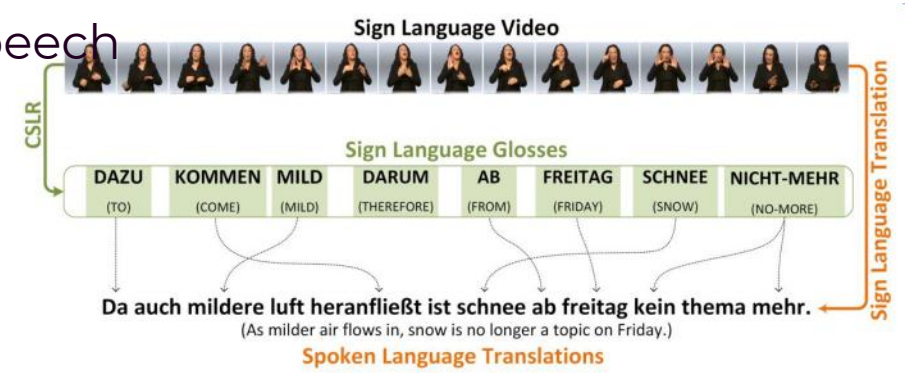
Sign Language Machine Translation

Translation directions



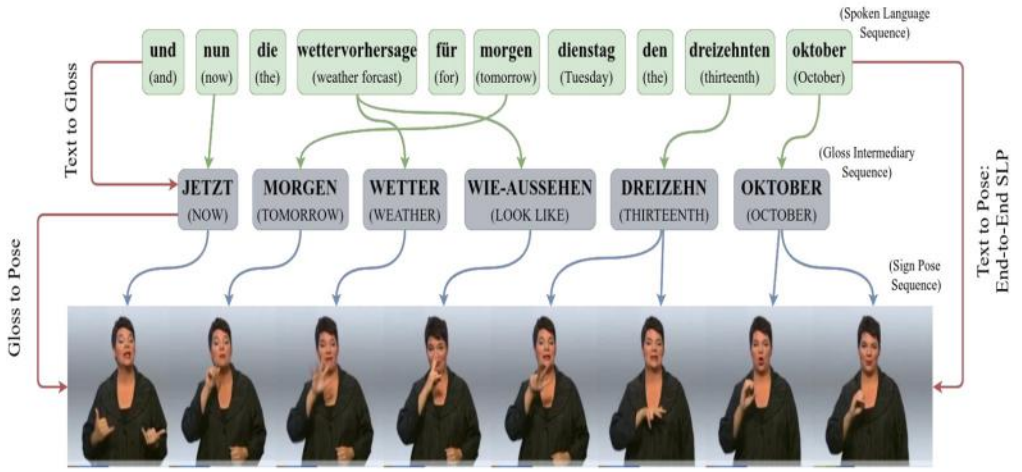
- **(Speech-to-)text to sign**
  - State-of-the-art machine translation
    - Avatar presenting signed MT output

- **Sign to text(-to-speech)**
  - Robust data-driven video recognition
    - State-of-the-art machine translation
      - Output in text/speech



What is sign language MT?

Two directions: **sign to spoken** or **spoken to sign**



Translation: NOT a simple sign to word mapping

To **learn this relationship** we need either:

- **Rules** that allow us to convert from sign to spoken or vice versa,
- **Data** so that the machine can learn these rules itself.

However, the data normally needs **curation**.

- Careful **alignment** and expensive annotation.
- We have many orders of magnitude **less data than we need**.

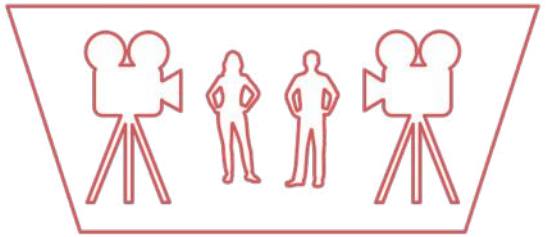
Manual annotation of gloss & recognition technology is therefore important to:

- automate the annotation process,
- provide alignment and tools to curate data automatically,
- condition the translation models.

Data (quality & quantity)

- Multidimensionality and multimodality of sign languages
  - Annotation
  - Data processing
- Data scarcity
  - Datasets of various types (broadcast, linguistic corpora, personal videos)
  - Language pairs
  - Approaches for translation (statistical vs lexical vs neural vs rule based)

Linguistic corpora



High quality

Variety of elicitation tasks

Source: Sign Language

Semi-spontaneous language production

Rich linguistic annotation + translations

Broadcasting data

Large quantity

News domain

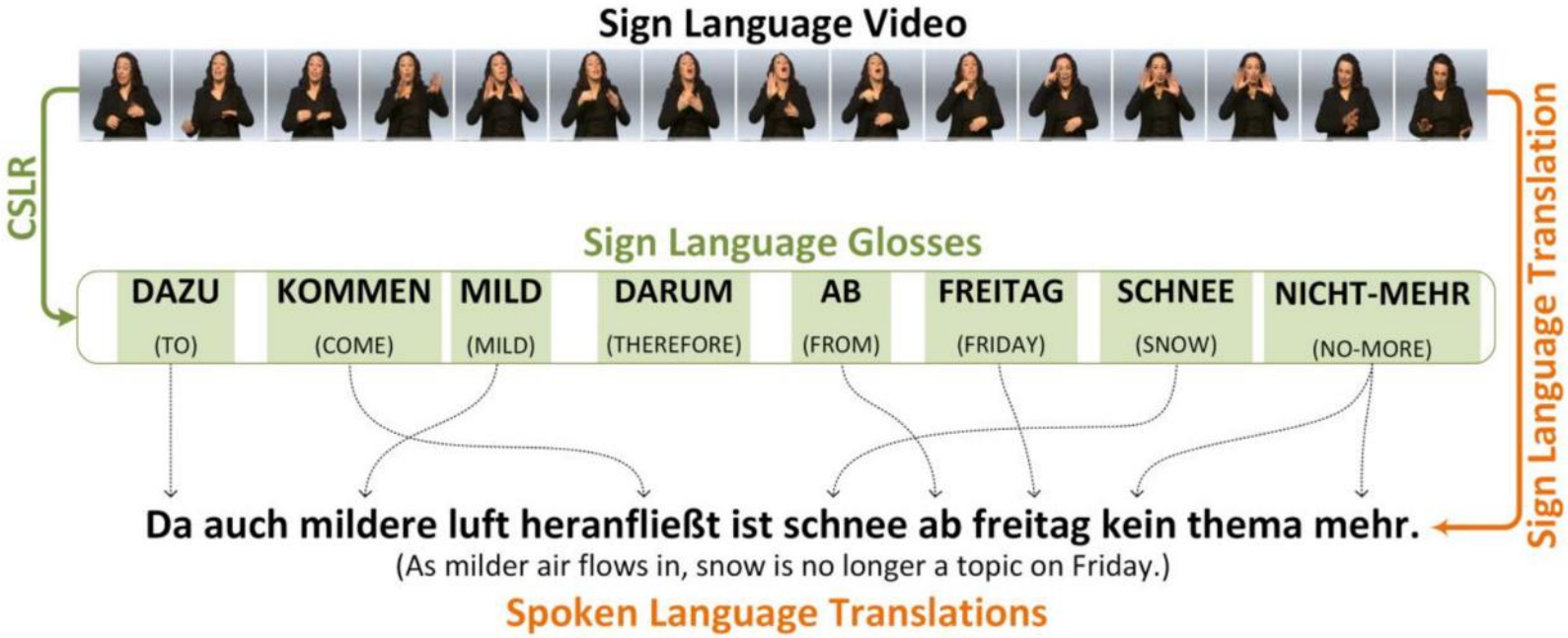
Source: spoken language

Interpreting under time pressure

Subtitles/Captions

Sign Language Recognition

Sign language recognition for translation



- CSLR: **C**ontinuous **S**ign **L**anguage **R**ecognition
- We use “Recognition” to refer to identifying specific signs within a continuous video sequence.
- Recognizing the sign is similar to “glossing” the data, which is what a linguist would do to annotate a sign language video.
- However, **translation means converting the underlying message into the equivalent spoken language sentence** (sign->spoken).

MT Output Presentation via Synthetic Signing (avatar representation)

Typical approach to SL representation

To form a sign/signed phrase:
Retrieve “motion plans” from a lexicon

- Mocap
- Traditional key frame
- Linguistic description
- Procedural
- Annotated gloss stream

And concatenate

But are all avatars appropriate for MT output presentation?

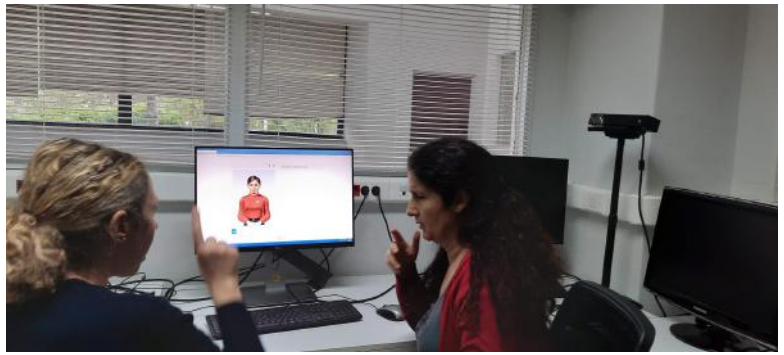
and then…

What about readability & user acceptance?

User experience

- Is it legible?
- Is it grammatical?
- Is it easy to read?

User Evaluation!



Technology for dynamic synthetic signing

- Avatars need to know context
- Avatars must support co-occurring events
- and they must be co-developed with their user communities

Current State in Data & Technologies

DATA or TECHNOLOGIES?

- Current machine learning approaches not applicable for SL processing (lack of data)
- Minority languages with already limited resources remain disconnected from technology
- Under-resourced languages have no chance to catch up
- Sign language processing technologies remain in the proof-of-concept state

⇨ **LLMs for SLs!**

easier

intElligent Automatic
SIgn languagE tRanslation

THANKS

EASIER project is funded by the EU's Horizon2020 Programme under Grant Agreement number 101016982

PROJECT-EASIER.EU

@EASIERPROJECT