

# Semi-Automatic Subtitle Alignment for Sign Languages

Maren Brumm, Sam Bigeard

Institute of German Sign Language and Communication of the Deaf, University of Hamburg, Germany



#### Problem:

- To provide subtitles for videos in sign language, the translation has to be aligned with the signing
- Doing this manually is time consuming but fully automatic alignment is often not accurate enough.

#### Our Proposal:

- Use automatically generated landmarks to aid human subtitle alignment
- Program generates an alignment proposal based on these landmarks
- User can change alignment directly but also alter, add or delete landmarks
- Program generates new alignment proposal based on the changes
- After a few feedback cycles the alignment is correct

## **Automatic Alignment**



Video





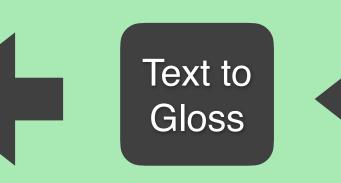


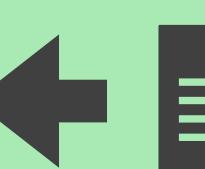


Confidence Gloss Frame SPÄTER10 0.7 SPÄTER10 0.8 SPÄTER10 117 0.4 **INDEX1** 118 0.3 119 ALT5A 0.6 ALT5A 0.9 120

Subtitle Gloss Confidence **ALTERN1** Er war dann 0.9 ein alter ALT2A 8.0 Kater. INDEX1 1.0 Er war 18 Jahre alt. INDEX4 0.9 NUM 1.0 JAHR1B 0.9

Possible glosses





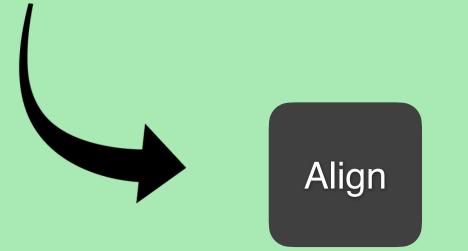
Subtitles

**Super Spotter** 

Proposes glosses from video

- Neural network based approach
- Detects approx. 2300 different types
- Available for BSL and DGS
- Developed and kindly shared by Ozge Mercanoğlu Sincan of Surrey University

## Spottings



## **Alignment Algorithm**

- 1. Filter spottings by confidence
- 2. Match spottings and possible glosses in sentences with altered Needleman-Wunsch algorithm
- 3. Place sentences according to matchings
- 4. If no match is found place sentence in gap
- 5. Adjust sentence length if longer or shorter than min/max length considering text length

Accuracy: median displacement of 1.8 sec

#### **Text to Gloss**

Proposes glosses from subtitles

- Word by word
- Dictionary approach with manual rules
- For special cases
- "sein" always ignored
- named entities map to FINGERSPELL + specific gloss if available
- Group glosses by sentences to avoid word order issues

## **Manual Alignment**

#### **General Functionality**

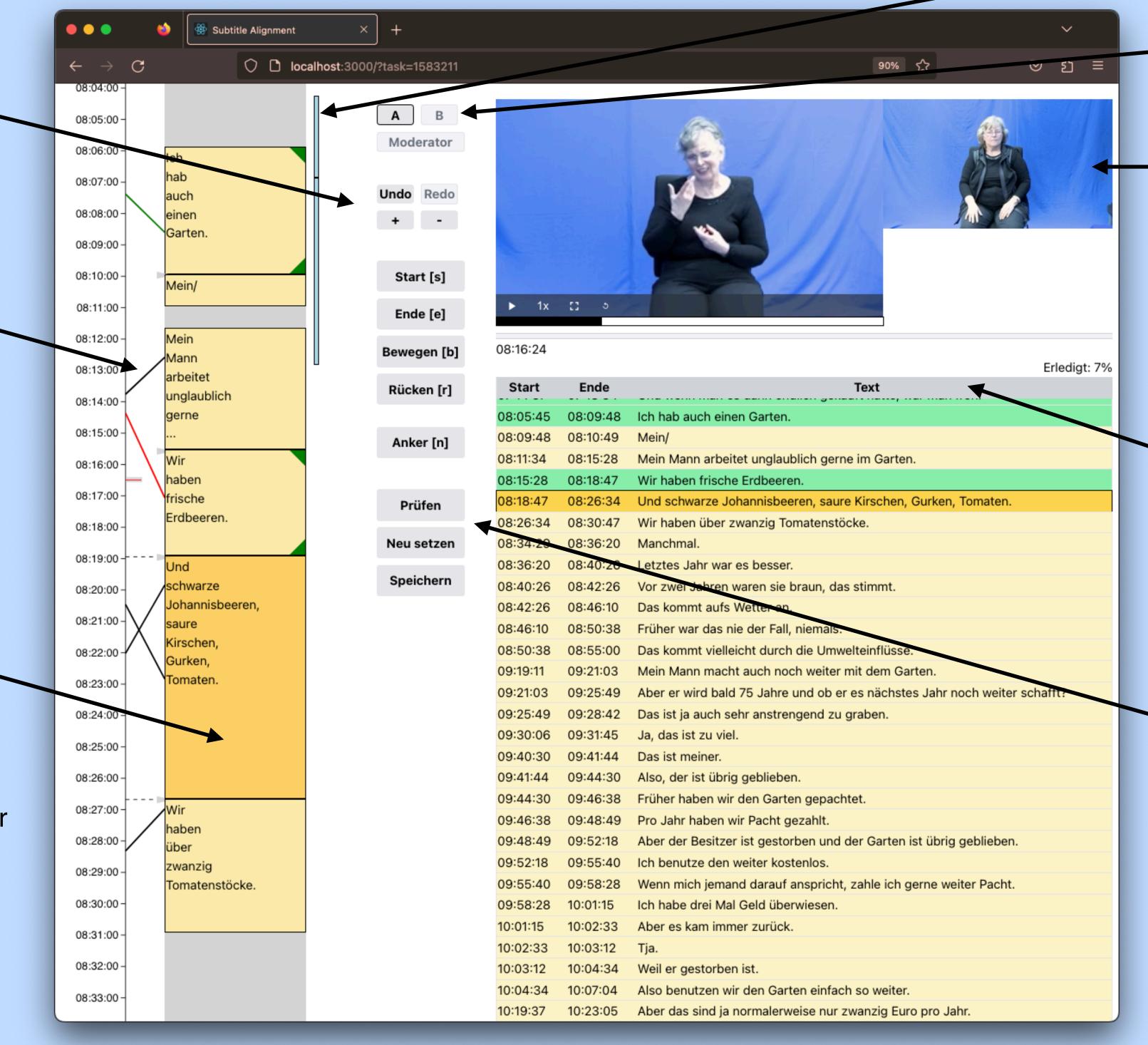
- Undo / Redo changes
- Zoom in and out of timeline
- Short keys for most operations

#### Landmarks

- Represented by lines
- Can be attached to words or gaps between subtitles (gray areas)
- Can alter time or word
- Black: unchecked
- Green: checked Red: checked but contradiction

#### **Subtitles**

- Move subtitle: alter start, end or complete position
- Optionally push down following subtitles so nothing overlaps
- Can split and join subtitles and alter text
- Green triangles: checked
- MultiSelect possible



#### **Second Signer**

- Shows in the timeline when the conversation partner is signing as well.
- Can switch between working on either signer in a conversation

#### Video

- Show active signer and second signer or moderator
- Play whole transcript or selected subtitles
- Video size and speed adjustable

#### **Table**

- For overview and navigation
- Green background: position checked
- Can mark everything up to current subtitle as correct

#### **Invoke Automatic**

- Check for contradictions in positioning
- New alignment proposal based on changes

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.

Poster presented at the Autumn School: Sign language data meets data science – data science meets sign linguistics at the IDGS Hamburg, Germany. September 25, 2023.

AKADEMIE DER WISSENSCHAFTEN IN HAMBURG

