



# Swedish Sign Language Corpus Ongoing Project: Annotation questions

Johanna Mesch, PhD & Lars Wallin, PhD  
Department of Linguistics, Stockholm University

Workshop "Sign Language Corpora: Linguistic Issues"  
University College London on 24-25 July, 2009

## Short introduction

- The Swedish Sign Language Corpus project (2009-2011)
- Includes recordings, documentation of sign language materials, annotations and transcriptions of signed materials from deaf native or near-native users of Swedish Sign Language

## Annotation conventions for corpus work

### Preparatory work

- Documentation and annotation of sign language materials
- Annotation conventions for Swedish SL

### Open access

- Backend – varied annotations
- Frontend – limited annotations

## Corpus, annotation, .....?

- The corpus consists of videofilms and fully glossed sign language text/discourse in the gloss tiers
- Annotation includes all kinds of notations of a sign language discourse which includes i.e. both transcription and tagging
- Transcription means sign language discourse rendered in writing
- Tagging means marking a sign or constituent with linguistic category concepts

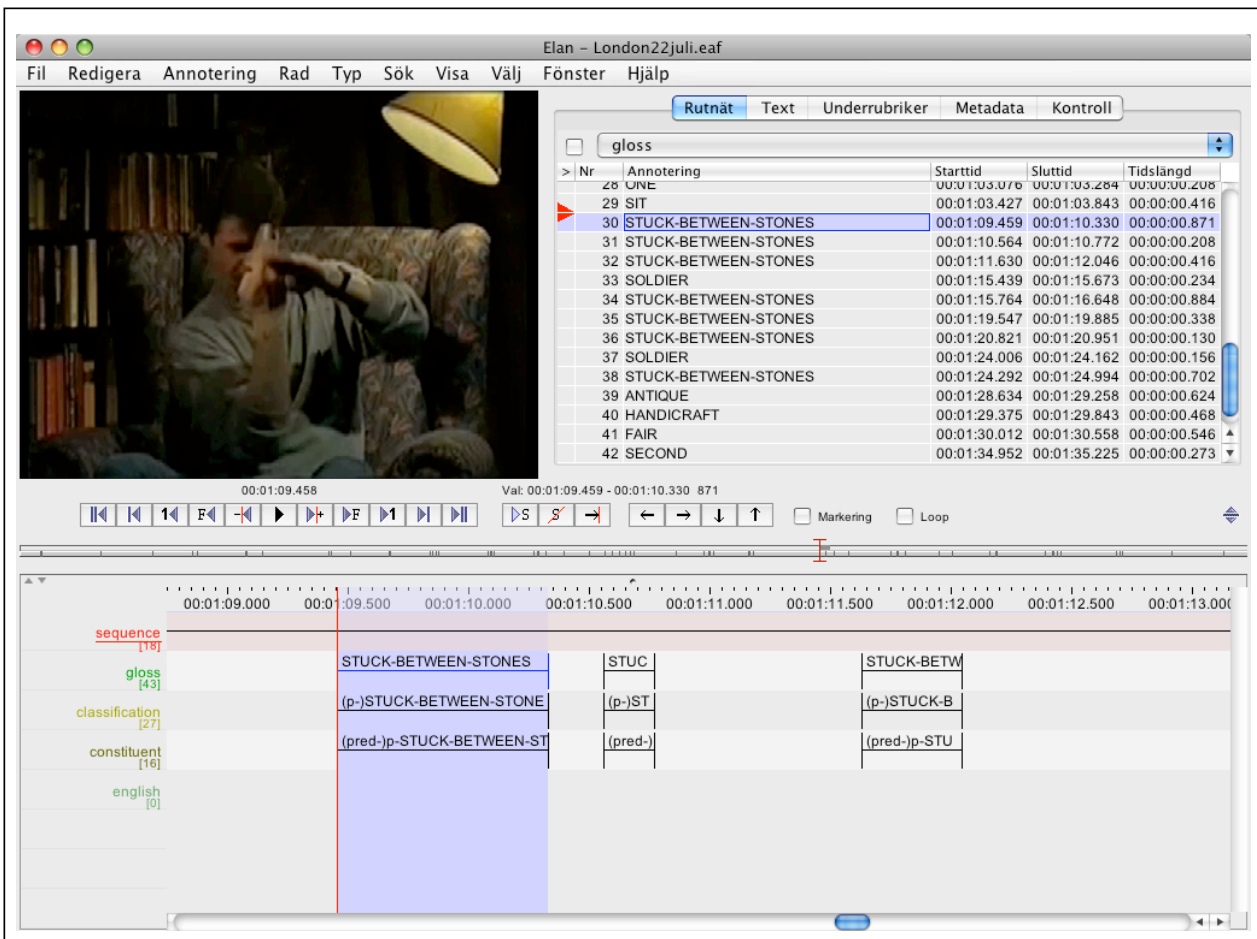
## Annotation issues

- All signs in a text are treated as lexical in the gloss tiers
- This included
  - a) Gestures
  - b) Fingerspellings
  - c) Polymorphemic signs

(Video show)

09-08-24 Mesch & Wallin; Department of Linguistics

5



Elan - London22juli.eaf

Fil Redigera Annotering Rad Typ Sök Visa Vålj Fönster Hjälp

Rutnät Text Underrubriker Metadata Kontroll

gloss

Nr	Annotering	Starttid	Sluttid	Tidslängd
28	ONE	00:01:03.076	00:01:03.264	00:00:00.208
29	SIT	00:01:03.427	00:01:03.843	00:00:00.416
30	STUCK-BETWEEN-STONES	00:01:09.459	00:01:10.330	00:00:00.871
31	STUCK-BETWEEN-STONES	00:01:10.564	00:01:10.772	00:00:00.208
32	STUCK-BETWEEN-STONES	00:01:11.630	00:01:12.046	00:00:00.416
33	SOLDIER	00:01:15.439	00:01:15.673	00:00:00.234
34	STUCK-BETWEEN-STONES	00:01:15.764	00:01:16.648	00:00:00.884
35	STUCK-BETWEEN-STONES	00:01:19.547	00:01:19.885	00:00:00.338
36	STUCK-BETWEEN-STONES	00:01:20.821	00:01:20.951	00:00:00.130
37	SOLDIER	00:01:24.006	00:01:24.162	00:00:00.156
38	STUCK-BETWEEN-STONES	00:01:24.292	00:01:24.994	00:00:00.702
39	ANTIQUE	00:01:28.634	00:01:29.258	00:00:00.624
40	HANDICRAFT	00:01:29.375	00:01:29.843	00:00:00.468
41	FAIR	00:01:30.012	00:01:30.558	00:00:00.546
42	SECOND	00:01:34.952	00:01:35.225	00:00:00.273

00:01:09.458 Vat: 00:01:09.459 - 00:01:10.330 871

sequence [18]  
gloss [43]  
classification [27]  
constituent [16]  
english [0]

sequence	gloss	classification	constituent	english
	STUCK-BETWEEN-STONES	(p-)STUCK-BETWEEN-STONE	(pred-)p-STUCK-BETWEEN-ST	
	STUC	(p-)ST	(pred-)	
	STUCK-BETW	(p-)STUCK-B	(pred-)p-STU	

## Annotation issues

All signs in the gloss tiers are written "bare", i.e. without classification tags like fs- or g-.

Such classification tags are written on other tiers instead.

## For example

Gloss tier: ANTIQUE HANDICRAFT FAIR

Class. tier: (cp-)HANDICRAFT\_FAIR

(Video show)

## Classification tags

- cp-HANDICRAFT\_FAIR compound
- fs-MONACO fingerspelling
- g-ATTRACT-ATTENTION gesture
- ncl-PERSON noun classifier
- vcl-SIT verbal classifier
- p-STUCK-BETWEEN-STONES polymorphemic

## Tack! Thank you!

Svensk teckenspråkskorpus  
Swedish Sign Language Corpus

[www.ling.su.se/tsp](http://www.ling.su.se/tsp)

