**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)
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