

The Department of Linguistics at Stockholm University participates in Swe-Clarín through two of its sections, Computational Linguistics and Sign Language.

The Section for Computational Linguistics carries out research related primarily to first-language acquisition and multilingual word alignment; for more information, see <http://www.ling.su.se/english/research/research-areas/research-in-computational-linguistics>. The section makes available several corpora and tools of potential interest to Swe-Clarín, including Stagger (a part-of-speech tagger with models for Swedish and Icelandic), the Stockholm Internet Corpus (SIC), and SUC-CORE (a subset of SUC annotated for coreference relations between noun phrases). An earlier infrastructural achievement was the Stockholm–Umeå Corpus (SUC), containing one million words with manually corrected part-of-speech tags, whose first version was developed 1989–97. Distribution of SUC was outsourced to Språkbanken in 2008, but the corpus is still being maintained within the department, with the latest version (SUC 3.0) having been released in 2012. For more information about the resources, see <http://www.ling.su.se/english/nlp>.



Much of the research in the Section for Sign Language is devoted to lexicography and corpora, particularly involving Swedish Sign Language (SSL). A breakthrough in the representation of sign-language data occurred in the mid-1990s when playback of video on personal computers became possible, replacing videotapes. Furthermore, the advent of systems for fine-grained, searchable annotation of video, such as ELAN (<https://tla.mpi.nl/tools/tla-tools/elan/>), has revolutionised the way in which sign language corpora can be represented and analysed. Work on the SSL lexicon started in 1988, and it is now comprising about 15,000 signs. Work on the SSL corpus started in 2003; it currently includes about 24 hours of annotated video from 42 signers of semi-spontaneous dialogues as well as monologues with narratives and elicitation tasks. Annotations include sign glosses and utterance-level translations into Swedish (see picture). Other resources include a corpus of tactile signing of deaf-blind signers and a learner corpus of SSL. For more information about the resources, see <http://www.ling.su.se/teckenspråksresurser>.

Last of the Summer Wine

Most of the national coordination team is back: Caspar has returned from his boat trip on the Danube, Stefan has come home from his boat trip on Disko Bay, and Lars has crisscrossed his way back from the fiords of northern Scandinavia. Nina will rejoin us by the end of September. Since the previous newsletter, the Swedish Language Bank has turned 40 (celebrated with informative as well as tasty festivities), Swe-Clarín has had a virtual meeting, and the UK has joined CLARIN as observer.

The autumn promises to bring exciting CLARIN-activities both nationally and internationally (see calendar below). We in the coordination team look forward to seeing many of you soon!



Winter is coming ...

Calendar

- 5–6 October: Nordic Clarin Network Workshop in connection to the Language Bank's autumn workshop on historical resources.
- 7–8 October: the coordination team visits Swe-Clarín-centres in Stockholm and Uppsala
- 9 October: workshop on digitalisation arranged by Digisam and the National Library of Sweden.
- 11 November: SND's autumn workshop under the theme "New Conditions for Research".
- 16–17 November: Workshop and Swe-Clarín partner meeting in Stockholm.
- 19–20 November: Meeting of the CLARIN ERIC General Assembly in Copenhagen.
- 11 December: virtual meeting for the Swe-Clarín partners, 10-12 am.

Partners

Swe-Clarín has nine partners from Lund, Gothenburg, Linköping, Stockholm and Uppsala, at universities and public authorities. A list and description of all partners may be found here:

<http://sweclarin.se/swe/centrum>

News

We will not go on spamming you. Should you want more info on Swe-Clarín, please sign up for the news list here:

http://lists.sweclarin.se/mailman/listinfo/news_lists_sweclarin.se

