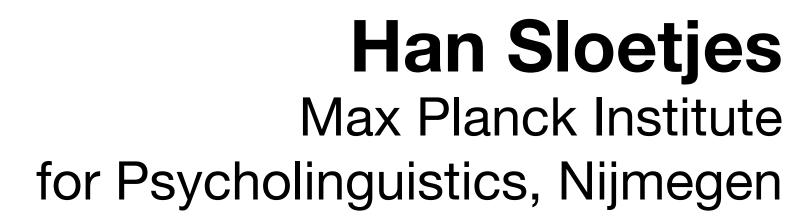
Using ELAN for annotating sign language corpora in a team setting



Onno Crasborn

Centre for Language Studies Radboud University Nijmegen



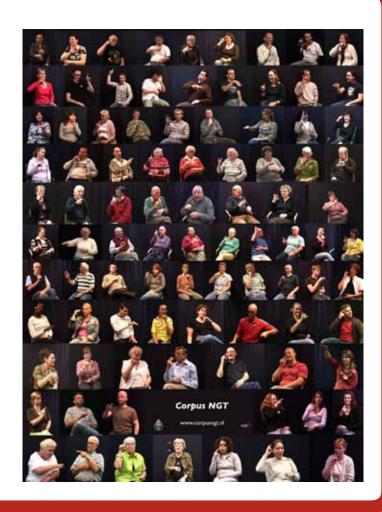


Background

ELAN was originally created with the idea of a single user working on a single document. Within the Corpus NGT project and subsequent linguistic research projects, the functionality to use ELAN to work on large numbers of documents in a research team with various interests has been slowly expanding. This poster shows some of these functions within ELAN version 3.8.1, and it shows how a complementary solution that has been developed to share files wihin a research group.

Context

Corpus NGT: 2375 files, used by 14 researchers and assistants simultaneously for different research projects. Total of >70 tiers.



help to visually group information in a large set of tiers.

Colours of tier labels Obs_Johan

- GlosCorrectie

HandvormL S1

HandvormR S1

LocatieR S1

SS_Mouth S1

HandvormL S2

- HandvormR S2

LocatieR S2

SignSpeakZin S1

SignSpeakZin S2

NDH SimConstr

NDH spreiding

NDH DomRey

NDH spr_loc

NDH spr_hv

NDH spr_ori

NDH betekenis

NDH rust

NDH echo

TurnTaking S1

TurnTaking S2

Mond_citatie S*

- Mond_type S1

- Mond_spr S1

Mond_citatie S2

- Mond_type S2

Mond_spr S2

Obs Yassine

Obs Onno

Obs_Els

Obs Anna

Obs_Inge

- Obs_Lianne

Obs Martine

Obs_Merel

Obs_Wim

Obs_Ellen

Obs_Sven

Obs_Richard

SS_zinsopmerkin

SS_Head mov S1

SS_Head mov S2

Eye blink S1

- Eye blink S2

Eye gaze S1

- Eye gaze S2

Body position S1

Body position S2

- Eye brow S1

- Eye brow S2

- Nose S1

Nose S2

Body mov S1

Body mov S2

Head position S1

Head position S2

Mond S2

Mond S1

Define search domain

mouth_dep

remarks

GlosR S2

GlosR S1

GlosCorrectie

Obs_Johan

Mond_citatie S2 Mond_citatie S1

- LocatieL S2

GlosL S2

GlosR S2

Tolk S1

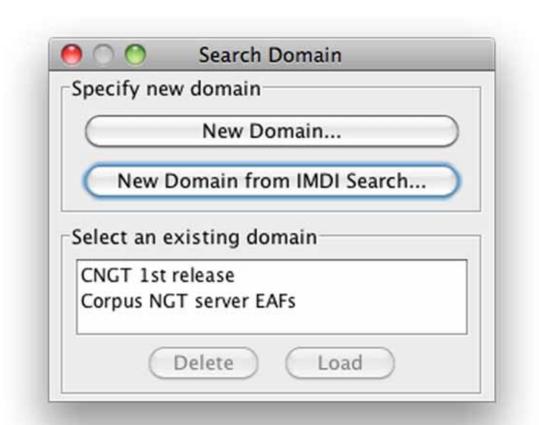
Tolk S2

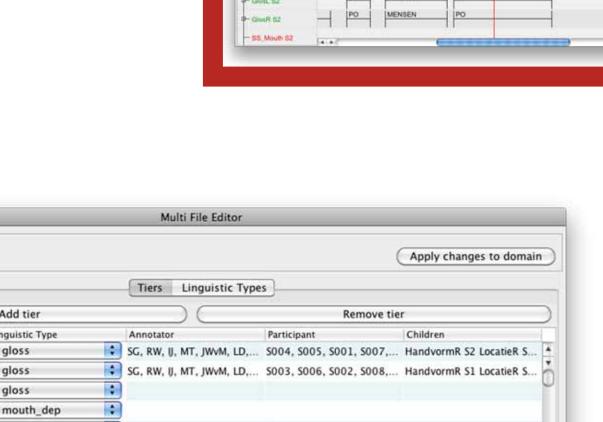
LocatieL S1

GlosR S1

GlosL S1

Search results from the IMDI metadata browser ('all men from Amsterdam') can be exported as a domain for ELAN searches.

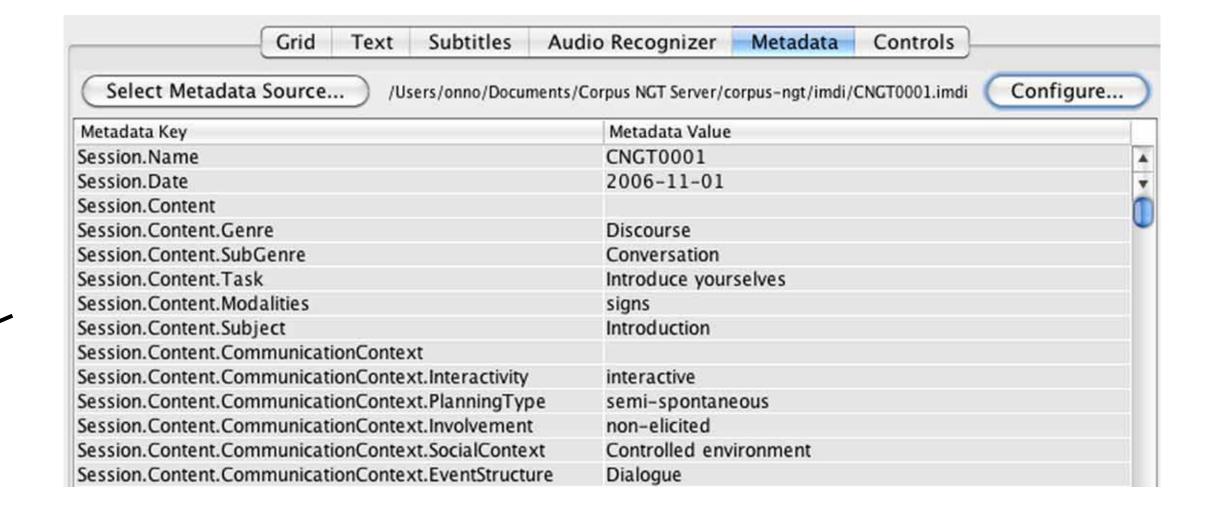




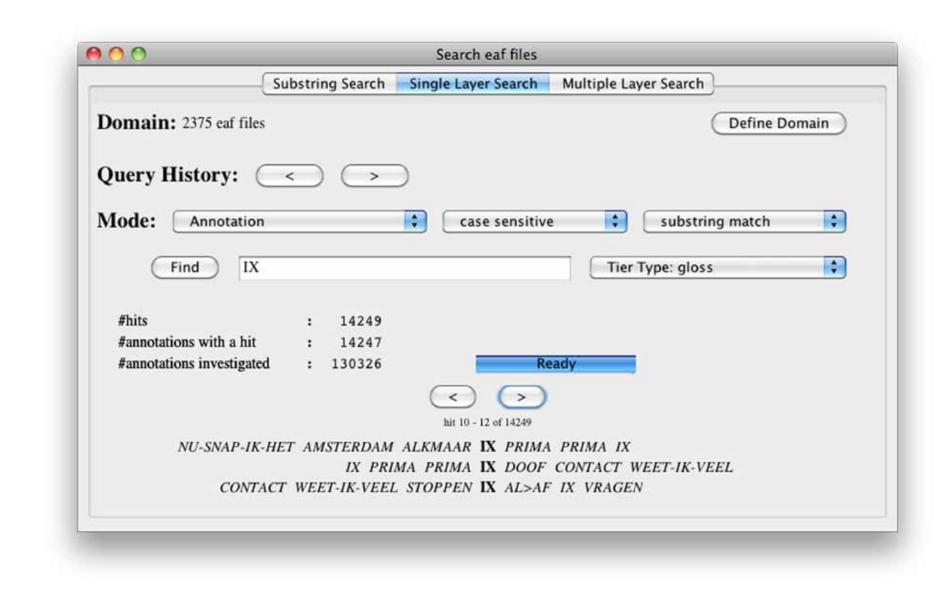
Close

The Multiple File Editor allows for adding and deleting tiers in a selected domain. By displaying all the properties that tiers have in this whole domain (Linguistic Type, Annotator, Participant), the function can help to keep large corpora well-organised.

Metadata from related IMDI documents can be visualised in ELAN.



Various actions can be performed on a selection of multiple files, including three types of searches.

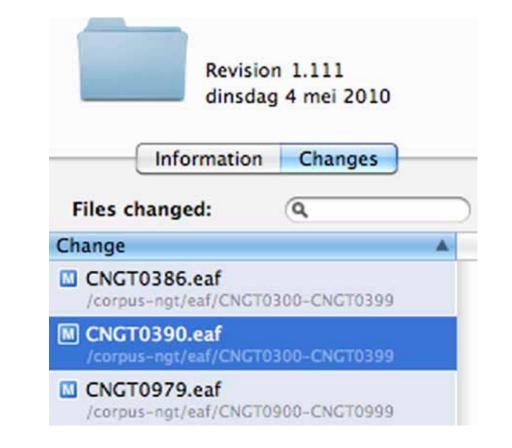


All annotation and IMDI documents are stored on a Subversion (SVN) server. Users work on a local copy of the archive, and publish their modifications to the server. This setup requires that users 'lock' the file they want to work on, and always ensure that they keep their local copies up to date.

▼ i corpus-ngt	7 19 april 2010	11:49	-	1.017	annedemeijer
▶ administratie	7 Today	11:55		1.109	johanros
▶ 🛅 eaf	7 Yesterday	11:50		1.093	ingezwitserlood
► 🛅 imdi	13 oktober 2009	12:56		202	o.crasborn
▶ ☐ PFSX_alleen_glossen	Yesterday	14:22	-	1.094	michahulsbosch
▶ m PFSX_Anna	31 januari 2010	12:52		536	annasafar
► PFSX_AnnaRU	1 februari 2010	11:40		540	annasafar
► PFSX_Annika	28 januari 2010	21:45		522	onno
► PFSX_Ellen	12 oktober 2009	09:59	-	187	o.crasborn
▶ PFSX_Els	12 oktober 2009	09:38		186	o.crasborn
▶ PFSX_fonologie	Yesterday	14:28		1.096	michahulsbosch

Different versions of preferences files have been created so that users see a selection of tiers that is relevant fo their work. The folder to use can be set in the ELAN preferences. Other actions on multiple files:

- find and replace
- generation of statistics
- create new documents based on template and sets of media files
- annotation 'scrubbing' (removal of superfluous spaces, tabs, and newline characters)
- export as word list for a selection of tiers



The SVN server keeps log info about who did what to which file. It stores a copy of every change ever submitted to the archive, and provides tools for seeing changes that have been made and resolving conflicts.