

ELAN

Multimedia Annotation Tool

Max-Planck-Institute for Psycholinguistics

<http://www.lat-mpi.eu/tools/elan>



ELAN

- written in Java programming language
- stores transcriptions in XML format (.eaf)
- available for Windows, Mac OS X, Linux
- sources available for non commercial use
- current version 3.4.0



Main window

The screenshot shows the Elan software interface with several components labeled:

- video player**: A window showing a video of people in a traditional setting.
- controls**: A set of playback controls including play, stop, and volume buttons.
- viewers & controls**: A menu bar with options like Grid, Text, Subtitles, and Controls.
- waveform viewer**: A visualization of the audio waveform.
- crosshair**: A vertical red line indicating the current time position.
- annotations**: A table of annotations with columns for Nr, Annotation, Begin Time, End Time, and Duration.
- timeline viewer**: A multi-layered timeline showing different tiers of data.
- tiers**: A list of tiers on the left side of the timeline viewer.

Nr	Annotation	Begin Time	End Time	Duration
1	he's speaking into his 1/2 bush knife	00:00:31.708	00:00:32.708	00:00:01.000
2	I am saying that he's talking into a half bushknife	00:00:37.141	00:00:38.501	00:00:01.360
3	everything we are saying is going inside	00:00:39.336	00:00:40.636	00:00:01.200
4	See this fellow: he seems to me like a little frog	00:00:45.426	00:00:47.106	00:00:01.680
5	it's collwecting all these things?	00:00:53.400	00:00:54.460	00:00:01.060
6	it's recording	00:00:55.521	00:00:56.056	00:00:00.535
7	You have sung them, sing them again	00:00:59.412	00:01:00.942	00:00:01.530
8	You sing those two parts of the Myââ mbwaa	00:01:01.120	00:01:02.462	00:00:01.342
9	You (P) ask him (Kp)	00:01:02.602	00:01:03.982	00:00:01.380
10	that thing, you sung those 2 parts of Myââ, sing them again	00:01:05.591	00:01:07.931	00:00:02.340
11	the next one	00:01:32.024	00:01:32.694	00:00:00.670

The timeline viewer shows the following tiers:

- Gloss1
- Gloss2
- Kp
- Pius
- K
- Ricky
- man outside vie
- child 1
- loudspeaker



Synchronized viewers

The screenshot shows a video player interface with a timeline from 00:00:17.000 to 00:00:26.000. The main track is 'Translation E' with the text: 'the shop, took the bone and took off as fast as he could. He ran far away up to the bridge. There he stopped to catch some breath, the bone in Water was streaming...'. Below it are tracks for 'Translation Du', 'Gloss RH English', 'Gloss LH English', 'Dir&Loc RH', 'Rep RH', and 'Gloss LH'. The 'Gloss RH English' track shows a sequence of actions: 'dog', 'CATCH', 'running d', 'dog disappears', 'BRIDGE', 'leg dog', 'BONE', 'leg dog', 'WATER', 'RIVER'.

timeline viewer

The screenshot shows an interlinear viewer with a grid of text. The top row is 'ref@unknown' with the text 'GOELONG-003'. Below it are several rows of text with annotations in blue and red boxes. The text includes: 'lap Ngen l@ nx{fn, Ngen nn@h@ ni t@ a mat g@ miskoom y...', 'lap Ngen l@ nx{fn Ngen nn@h@ ni t@ a mat g@ miskoom **', 'receive Ngen put inside Ngen this 3Sg.l FOC FOC women PREF elder **', and 'He...'. The grid has columns for different parts of the text and rows for different annotations.

interlinear viewer

The screenshot shows a grid viewer with a table of translation data. The table has columns: 'Nr', 'Annotation', 'Translation Dutch', 'Begin Time', 'End Time', and 'Duration'. The data is as follows:

Nr	Annotation	Translation Dutch	Begin Time	End Time	Duration
1	A dog was walking around. When he came across...	Een hond liep op zijn gemak. Toen hij bij een sl...	00:00:00.630	00:00:10.230	00:00:09.600
2	Hmm, that looked good.	Hij begon te watertanden: dat ziet er lekker uit!	00:00:10.230	00:00:11.990	00:00:01.760
3	He looked around carefully, nobody there.	Hij keek voorzichtig rond, niemand te zien.	00:00:12.010	00:00:15.640	00:00:03.630
4	He ran into the shop, took the bone and took off...	Hij rende snel de winkel in, pakte het bot en ren...	00:00:15.660	00:00:21.910	00:00:06.250
5	There he stopped to catch some breath, the bon...	Daar bleef hij staan om uit te hijgen met het bot i...	00:00:21.920	00:00:24.850	00:00:02.930
6	Water was streaming under the bridge.	Er stroomde water onder de brug door.	00:00:25.000	00:00:27.830	00:00:02.830
7	The dog, still panting, looked into the water	De hond keek in het water terwijl hij nog stond n...	00:00:27.850	00:00:31.180	00:00:03.330

grid viewer

text viewer

The screenshot shows a text viewer with a detailed view of the translation data. It includes a table with columns: 'Grid', 'Text', 'Subtitles', and 'Controls'. The data is as follows:

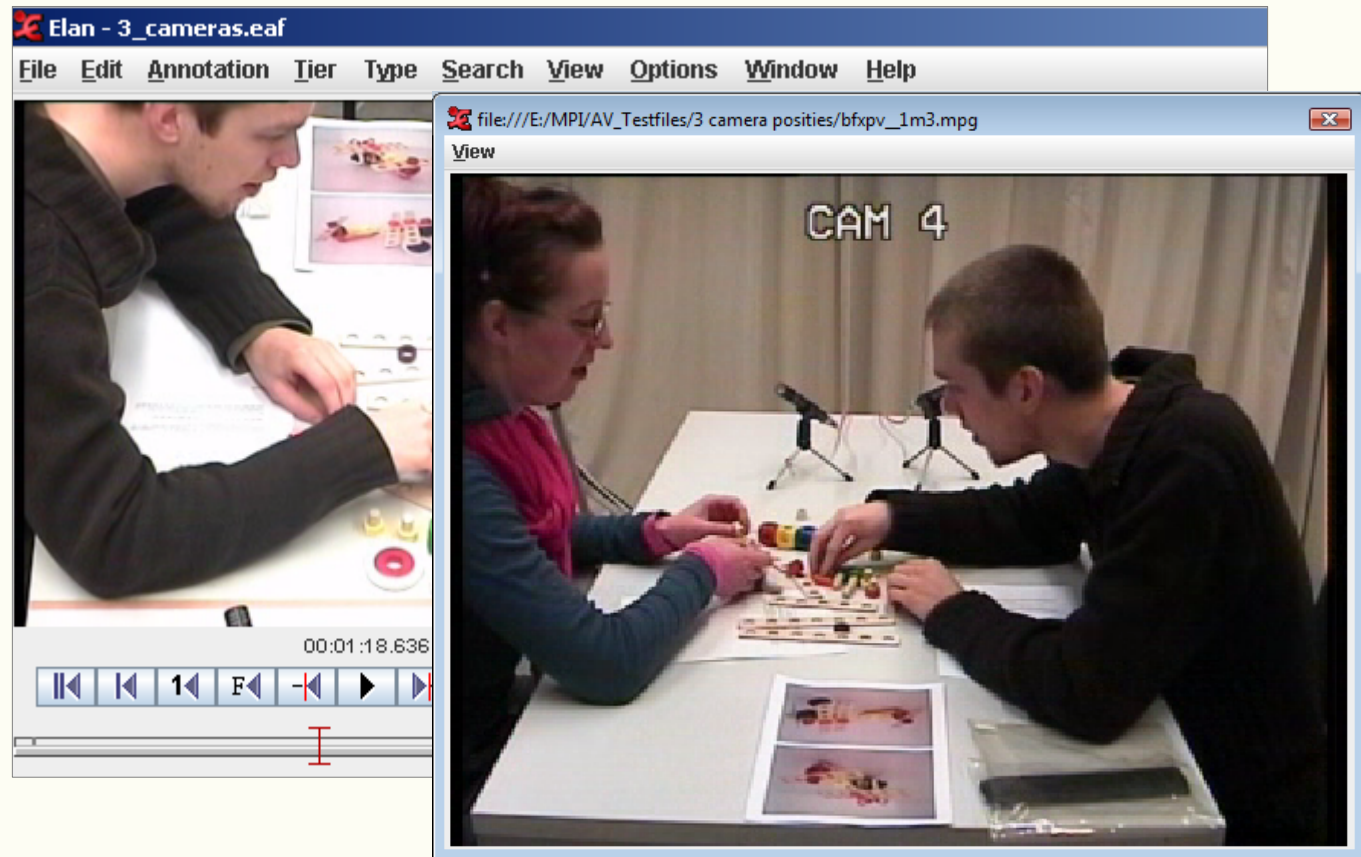
Grid	Text	Subtitles	Controls
Translation Dutch	Een hond liep op zijn gemak. Toen hij bij een slagerszaak kwam, zag hij een bot liggen. · Hij begon te watertanden: dat ziet er lekker uit! · Hij keek voorzichtig rond, niemand te zien. · Hij rende snel de winkel in, pakte het bot en rende er zo snel als hij kon mee weg. Hij rende ver weg tot aan de brug. · Daar bleef hij staan om uit te hijgen met het bot in zijn bek. · Er stroomde water onder de brug door. · De hond keek in het water terwijl hij nog stond na te hijgen · Opeens schok hij van wat hij zag, in het water was een andere hond. · De hond blafte. · De hond zag dat de andere hond ook een bot had. · Hij heeft een groot bot, de mijne is maar klein. · Ik wil dat bot van hem afpakken! · De hond gromde naar de hond in het water: Vechten? Kom op dan! · Hij gromde: "Ik wil het bot afpakken!" · Hij sperde zijn bek open. Plons! Het bot viel in het water en werd door de stroom meegevoerd. · Hij keek in het water, de andere hond was weg! · Pas toen de golven weer minder waren, zag hij dat het gezicht in het water van zichzelf was. · Oh nee, dat niet... · Het bot was weg gedreven, en nu was het verdwenen. · Wat stom van mij!		
Translation English	He ran into the shop, took the bone and took off as fast as he could. He ran far away up to the bridge.		
Gloss RH English			
Dir&Loc RH			

subtitle viewer



Multiple videos

- display of up to 4 videos
- videos are detachable



Synchronization of videos

- determine an offset per video

The screenshot displays the Elan software interface for video synchronization. The window title is "Elan - NGT_AH_fab5.eaf". The menu bar includes "File", "Edit", "Annotation", "Tier", "Type", "Search", "View", "Options", "Window", and "Help". The main area is split into two video panels: "CAM 2" on the left and "CAM 3" on the right, both showing a woman clapping. Below the panels, a central box labeled "offset" has two red arrows pointing to two red ovals. The left oval contains "Player 1 Offset: 00:00:07.369" and "NGT_AH_fab5_b.mpg". The right oval contains "Player 2 Offset: 00:00:07.350" and "NGT_AH_fab5_f.mpg". At the bottom, there are controls for "Offset" (radio buttons for "Use Absolute Offsets" and "Use Relative Offsets", and a button "Apply Current Offsets") and "Player" (radio buttons for "All", "Player 1", and "Player 2"). A video playback control bar is also visible.



The platform dependent part

- Windows:
 1. Windows Media Player (Direct Show, via native library)
 2. Quick Time (QT for Java, library by Apple)
 3. Java Media Framework (Sun Microsystems)
- Mac OS X:
 1. Quick Time (QT for Java)
- Linux:
 1. Java Media Framework

Supported file types depends on the framework



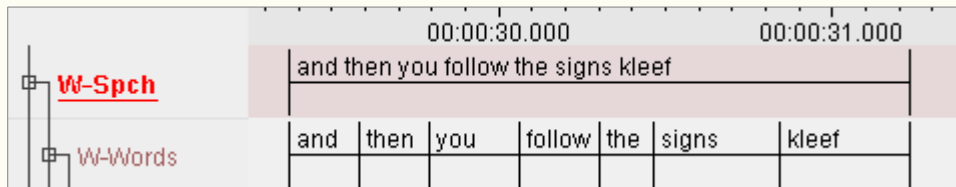
Key concepts

Tiers, linguistic types, annotations, controlled vocabularies

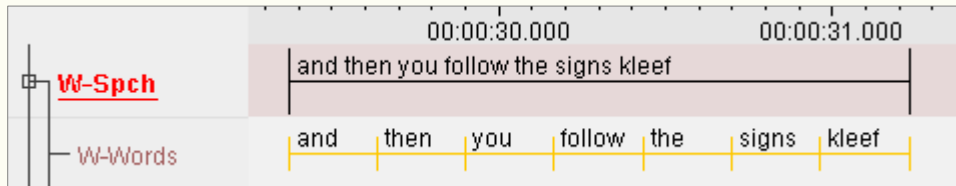
- a tier logically groups annotations that share the same constraints on structure, content and / or time alignment characteristics
- tiers can be part of a tier hierarchy, a tier can have a parent tier and dependent tiers
- an annotation document can contain any number of tiers and annotations
- a controlled vocabulary is a list of predefined annotation values on a tier



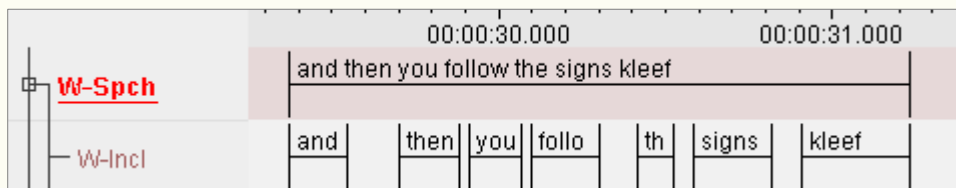
Inter-tier relations



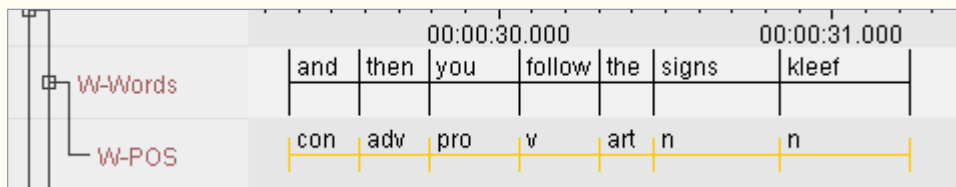
- "Time Subdivision"
Time aligned subdivision of parent interval, no gaps



- "Symbolic Subdivision"
Ordered, symbolic subdivision of parent interval



- "Included In"
Time aligned subdivision of parent interval, gaps are allowed



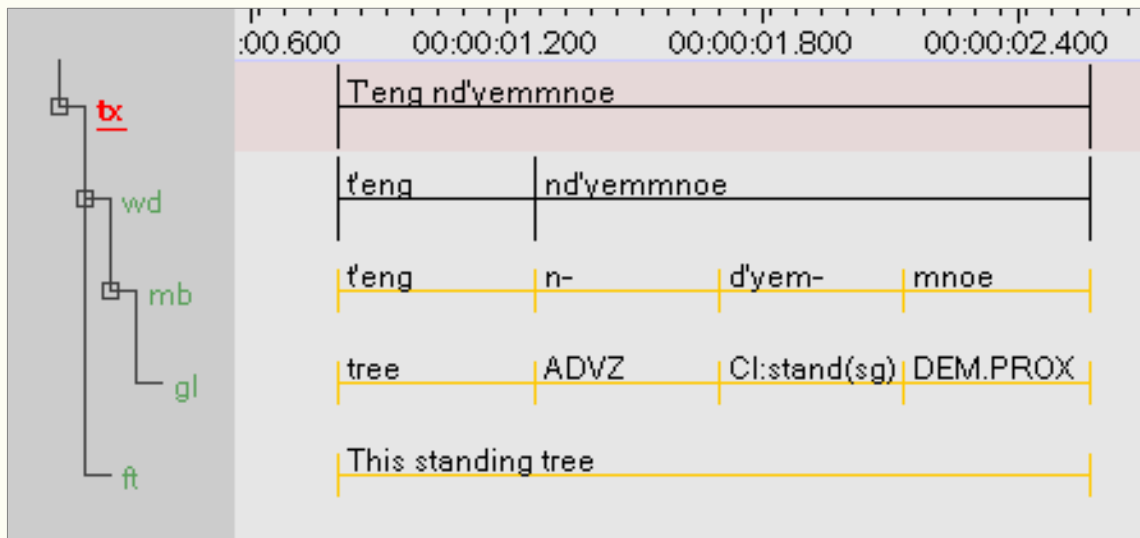
- "Symbolic Association"
One-to-one association with the parent annotation



Example transcription

tx (text)	<i>T'eng nd'yemmnoe.</i>			
wd (word)	<i>t'eng</i>	<i>nd'yemmnoe</i>		
mb (morpheme break)	<i>t'eng</i>	<i>n-</i>	<i>d'yem-</i>	<i>mnoe</i>
gl (gloss)	<i>tree</i>	<i>ADVZ</i>	<i>Cl:stand(sg)</i>	<i>DEM.PROX</i>
ft (free translation)	<i>This standing tree.</i>			

interlinear text example
(text editor, Shoebox)



same example in ELAN



Annotating

- manual creation of annotations
- templates documents
- input methods for a wide range of script systems

lookup list

virtual keyboard

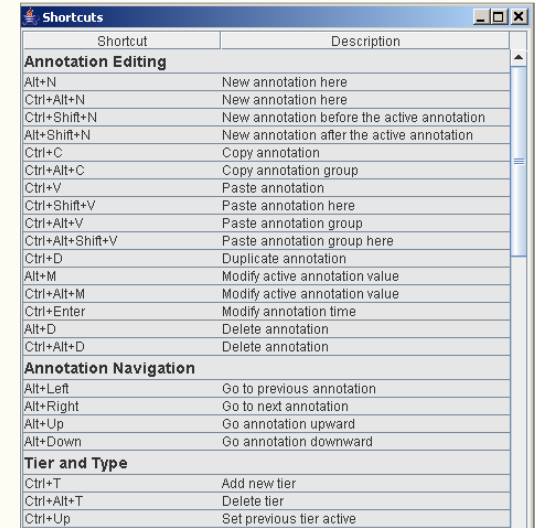


The screenshot displays a software interface for audio annotation. At the bottom, a 'Georgian (Heinecke) keyboard map' is shown, featuring a standard QWERTY layout with Georgian characters. Above it, a 'virtual keyboard' is indicated by a box. To the right, a 'Lookup list' window is open, displaying a vertical list of Georgian text: 無, 於, 五, 物, 務, 武, 誤, 鳴, 舞, 亞. The background shows an audio waveform with a red cursor and time markers (00:00:17.443 - 00:00:19.953).

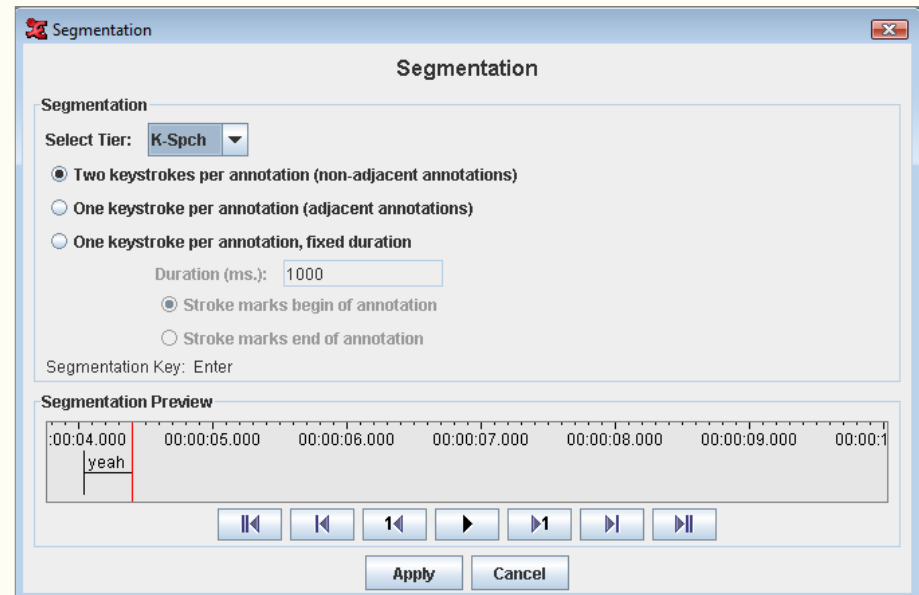


Editing

- multiple undo / redo
- copy, duplicate, paste annotations
- keyboard shortcuts
- special annotation creation modes
 - tokenizing
 - filtering
 - semi automatic segmentation
 - annotations from overlaps



Shortcut	Description
Annotation Editing	
Alt+N	New annotation here
Ctrl+Alt+N	New annotation here
Ctrl+Shift+N	New annotation before the active annotation
Alt+Shift+N	New annotation after the active annotation
Ctrl+C	Copy annotation
Ctrl+Alt+C	Copy annotation group
Ctrl+V	Paste annotation
Ctrl+Shift+V	Paste annotation here
Ctrl+Alt+V	Paste annotation group
Ctrl+Alt+Shift+V	Paste annotation group here
Ctrl+D	Duplicate annotation
Alt+M	Modify active annotation value
Ctrl+Alt+M	Modify active annotation value
Ctrl+Enter	Modify annotation time
Alt+D	Delete annotation
Ctrl+Alt+D	Delete annotation
Annotation Navigation	
Alt+Left	Go to previous annotation
Alt+Right	Go to next annotation
Alt+Up	Go annotation upward
Alt+Down	Go annotation downward
Tier and Type	
Ctrl+T	Add new tier
Ctrl+Alt+T	Delete tier
Ctrl+Up	Set previous tier active



Controlled Vocabularies

- user definable lists of possible annotation values

CV name

CV entries

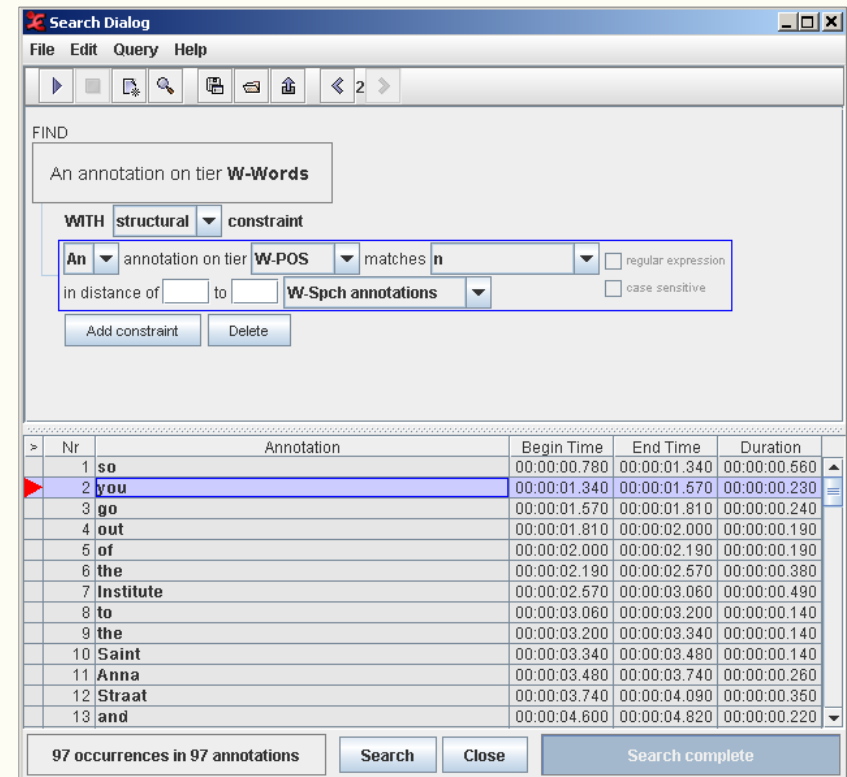
import a CV

select an entry from the list



Single file search

- create complex search queries
- combination of temporal and structural constraints
- use of regular expressions
- save and load queries
- export result to tab-delimited text
- search and replace



The screenshot shows a 'Search Dialog' window with a menu bar (File, Edit, Query, Help) and a toolbar. The search criteria are: 'An annotation on tier W-Words' WITH 'structural' constraint. The search matches 'n' in 'W-POS' annotations, with options for 'regular expression' and 'case sensitive'. The results table below shows 13 annotations with their respective begin and end times and durations.

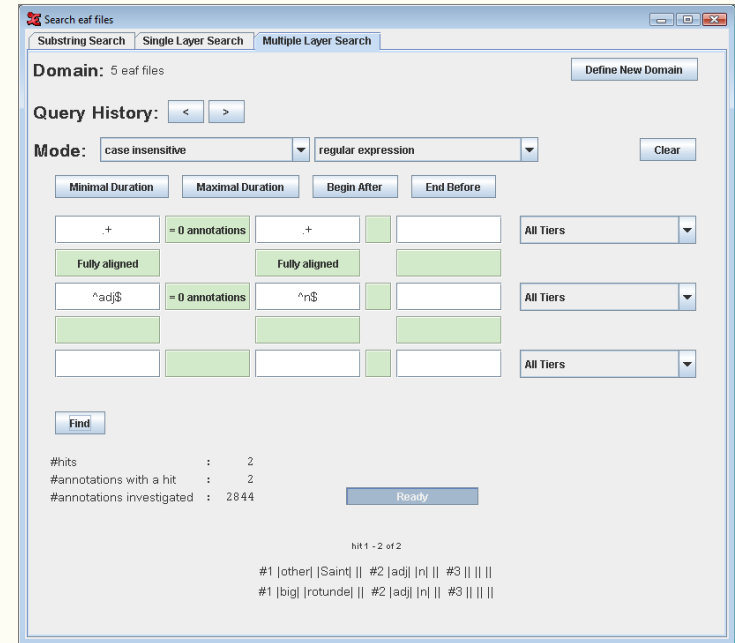
Nr	Annotation	Begin Time	End Time	Duration
1	so	00:00:00.780	00:00:01.340	00:00:00.560
2	you	00:00:01.340	00:00:01.570	00:00:00.230
3	go	00:00:01.570	00:00:01.810	00:00:00.240
4	out	00:00:01.810	00:00:02.000	00:00:00.190
5	of	00:00:02.000	00:00:02.190	00:00:00.190
6	the	00:00:02.190	00:00:02.570	00:00:00.380
7	Institute	00:00:02.570	00:00:03.060	00:00:00.490
8	to	00:00:03.060	00:00:03.200	00:00:00.140
9	the	00:00:03.200	00:00:03.340	00:00:00.140
10	Saint	00:00:03.340	00:00:03.480	00:00:00.140
11	Anna	00:00:03.480	00:00:03.740	00:00:00.260
12	Straat	00:00:03.740	00:00:04.090	00:00:00.350
13	and	00:00:04.600	00:00:04.820	00:00:00.220

97 occurrences in 97 annotations Search Close Search complete



Multiple file search

- add ELAN files to a 'basket'
- simple and structured search:
several matching types,
in selected set of tiers
- concordance and frequency
view
- export results to
tab-delimited text file
- open a file by a
click on a result



Search eaf files

Annotations

Search going [Search] [Define search domain] [Export] [Search complete]

regular expression case sensitive

23 occurrences in 23 annotations in 5 files (0.14 seconds)

Nr	File	Tier	Before	Annotation	After	Begin Time	End Time	Duration
1	elan-demo	W-RGMe		Going out of the Institute	Going out of MPI to...	00:00:00.780	00:00:04...	00:00:03...
2	elan-demo	W-RGMe	Going out of the...	Going out of MPI to Saint	Going along St. An...	00:00:04.600	00:00:10...	00:00:05...
3	elan-demo	W-RGMe	Going out of M...	Going along St. Anna Street	Going along St. An...	00:00:31.320	00:00:31...	00:00:00...
4	elan-demo	W-RGMe	Going along St...	Going along St. Anna Street	A roundabout	00:00:32.379	00:00:34...	00:00:01...
5	elan-demo	W-RGMe	A roundabout	Going to the direction of	Going along Oranj...	00:00:01.570	00:00:01...	00:00:00...
6	elan-demo	W-RGMe	Going to the dir...	Going along Oranje Singel	Going to the direct...	00:00:05.370	00:00:05...	00:00:00...
7	elan-demo	W-RGMe	Going along Or...	Going to the direction of	Going down	00:00:31.320	00:00:31...	00:00:00...
8	elan-demo	W-RGMe	Going to the dir...	Going down	a roundabout	00:00:32.920	00:00:33...	00:00:00...
9	elan-demo	W-RGMe	a roundabout	Going down	Going down	00:00:00.630	00:00:10...	00:00:09...
10	elan-demo	W-RGMe	Going down	Going down	a roundabout	00:00:27.850	00:00:31...	00:00:03...
11	elan-demo	W-RGMe	a roundabout	Going down	Going down	00:00:31.370	00:00:33...	00:00:02...
12	elan-demo	W-RGMe	Going down	Going down	Going to Kleve	00:00:33.820	00:00:35...	00:00:01...
13	elan-demo	W-RGMe	Going down	Going to Kleve		00:00:35.280	00:00:38...	00:00:03...
14	rabbitENGO1_test	speech		(going) hundred and eighty	backwards and for...	00:00:43.490	00:00:48...	00:00:05...
15	rabbitENGO1_test	speech	turning around i...	it's going	but meandering fr...	00:00:58.850	00:01:01...	00:00:02...
16	rabbitENGO1_test	speech	but meanderin...	(going) forward to it	for a few steps	00:01:01.570	00:01:11...	00:00:09...
17	rabbitENGO1_test	speech	turning round b...	(going) in a kind of loop	It's coming down i...	00:00:00.840	00:00:01...	00:00:01...
18	rabbitENGO1_test	speech	but this time it's...	(going) round three	as it's moving rou...	00:00:01.920	00:00:04...	00:00:02...
19	rabbitENGO1_test	speech		(going) hundred and eighty	backwards and for...	00:00:10.430	00:00:11...	00:00:00...
20	rabbitENGO1_test	speech	turning around i...	it's going	but meandering fr...	00:00:15.870	00:00:16...	00:00:01...
21	rabbitENGO1_test	speech	but meanderin...	(going) forward to it	for a few steps	00:00:17.860	00:00:18...	00:00:00...
22	rabbitENGO1_test	speech	turning round b...	(going) in a kind of loop	It's coming down i...	00:00:18.710	00:00:20...	00:00:01...
23	rabbitENGO1_test	speech	but this time it's...	(going) round three	as it's moving rou...	00:00:21.890	00:00:22...	00:00:00...



Interlinear output

- customizable printing and export to text/html file
- selection and sorting of tiers
- selection of time code format
- line and 'block' wrapping

Set Layout Options

W-SpCh and then you go the other, Saint Anna Straat to this to the center of the town, to this big rotunde.
W-Words and then you go the other Saint Anna Straat to this to the center
W-POS con adv pro v art adj n n n prep dem prep art n
W-IPA and ðen ju: gau ðt ɔ̃t̃e sant ana stra t to ðas to ða senta af ða taun ju ðas big rotunde
TC 00:00:04.600 - 00:00:10.070

W-Words of the town to this big rotunde
W-POS prep art n prep dem adj n
K-SpCh ja
TC 00:00:10.060 - 00:00:10.460

W-SpCh and you follow then the sign Kleef that's the oranje single
W-Words and you follow then the sign Kleef that 's the Oranje Single
W-POS con pro v adv art n n dem v art n n
W-IPA and ju: fɛlaʊ ðan ða saan kle f ðaets ða oranja sanj
TC 00:00:11.190 - 00:00:13.240

W-SpCh then you follow the sign kleef you come down
W-Words then you follow the sign Kleef you come down
W-POS adv pro v art n n pro v adv
W-IPA ðen ju: fɛlaʊ ða saan kle f ju: kam daen
TC 00:00:15.330 - 00:00:17.190 00:00:17.570 - 00:00:18.190

K-SpCh ja
TC 00:00:14.620 - 00:00:14.990

W-SpCh you know eh after this trajanus plein you come down to the
W-Words you know eh after this Trajanus Plein you come down to the
W-POS pro v post prep dem n n pro v adv prep art
W-IPA ju: nɔ: a afta ðas trajanus ple an ju: kam daen to ða
TC 00:00:18.990 - 00:00:21.810 00:00:21.880 - 00:00:22.930

W-SpCh rhine eh valley yeah that's another eh
W-Words Rhine eh valley yeah that 's another eh
W-POS n post n int dem v adj part
W-IPA raan a vaelt jɛ: ðaets anaða a
TC 00:00:23.470 - 00:00:25.050 00:00:25.080 - 00:00:27.010

K-SpCh there is another plein
TC 00:00:22.830 - 00:00:23.870

W-SpCh kind of rotunde
W-Words kind of rotunde
W-POS n prep n
W-IPA

Tiers

<input type="checkbox"/>	W-IPA	12
<input checked="" type="checkbox"/>	K-SpCh	12
<input type="checkbox"/>	wts	12
<input type="checkbox"/>	wss	12
<input type="checkbox"/>	wsipa	12
<input type="checkbox"/>	W-RGU	12

Font Sizes

What

Show Selection Only

Show Tier Labels

Show Time Code hh:mm:ss.ms

Show Empty Slots

Hide Empty Lines

Show All Tiers

How

Width: 559

Height: 783

Wrap Blocks Within Block

Wrap Lines

Sort By Tier Hierarchy

Line Spacing: 0

Block Spacing 20

Apply Changes Page Setup Print

preview area

select and sort tiers

- what to include

- how to format



Import / export

- Import:
 - Shoebox/Toolbox (.txt)
 - Transcriber (.trs)
 - Chat (.cha)
 - Praat (.TextGrid)
 - CSV/Tab-delimited text (.csv)
- Export:
 - Shoebox/Toolbox
 - Chat
 - Tab-delimited text
 - Interlinear text
 - Praat
 - Word list, list of annotations



LAT Tools

- tools developed at the MPI for Psycholinguistics



Language Archiving Technology

Tools



Addit a tool to 'plug' notes to elements of the MPI archive.



AMS a tool to grant and deny the access to a (part of a) corpus.



Annex the Annotation Exploration tool in the MPI web-based framework for archive exploration and enrichment



Elan Multimedia Annotator



Indi create, search through and modify metadata for linguistic resources



Lamus Language Archive Management and Upload System



Lexus a web-based lexicon tool



Synpathy a tool for manual syntactical annotation



ELAN

<http://www.lat-mpi.eu/tools/elan>

