

# ELAN - recent and (possible) future developments

Han Sloetjes The Language Archive - Max Planck Institute for Psycholinguistics Nijmegen, The Netherlands





A tool for annotating (multiple) audio and/or video files

Produces XML (eXtensible Markup Language) files: EAF (ELAN Annotation File)

Is free software, the sources are available for download

Mainly written in the Java programming language

Runs on Windows, Mac OS and Linux\* (\* in principle..)

Current version 3.9.1

Website: http://www.lat-mpi.eu/tools/elan/





A tool that is used in various types of research:

language documentation projects
sign language research
gesture research
multimodal interaction research

A tool that allows for multi-layered annotation, the layers (tiers) can be part of a hierarchy

A tool in which the annotation process (segmentation and labeling) is mainly a manual labor





Manually annotating media is a very time consuming process

A lot of work has already been done in the field of pattern recognition in audio and/or video files

Integration of this type of software in annotation programs seems rare

Automatic segmentation by pattern recognition based software seems feasible

In **ELAN** a Silence Recognizer has been included since version 3.6 (Aug 2008)

## Semi-automatic annotation



ELAN 3.9.1 File Edit Annotation Tier Type Search View Options Window Help	🚺 🕕 📕 🕴 🎅 🖣 💻 🖅 (75%) zo 1
Elan - pear story.eaf	
Grid Text Subt	itles Audio Recognizer Video Recognizer Metadata Controls eaks of voiced timespans) File(s): Ø pear.wav
Selections Add Remove Add Tier	Parameters Pitch ceiling [Hz], eg. 300 male / 550 female (150.0 - 1500.0) 604.8 Intensity change [dB] to start/end a peak (0.5 - 5.0) 2.0 Minimum amplitude (01) for pitch analysis (0.01 - 0.3)
0::0:16.840  Segmentations  Create Tier(s)  0::0::0::00::00::00::00::00::00::00	Progress Ready Start Report
	Node 🗌 Loop Mode 🔶
#mtenskyc8 55.3015	MMMMM 38.9515 MMMMM
#pitch.Hz 340.3425	504.576
00 00:00:09.000 00:00:10.000 00:00:11.000 00:00:12.000 00:00:13.000 00:00:14.000	00:00:15.000 00:00:16.000 00:00:17.000 00:00:18.000 00:00:19.000 00:01
to 00:00:09.000 00:00:10.000 00:00:11.000 00:00:12.000 00:00:13.000 00:00:14.000 Event	00:00:15.000 00:00:16.000 00:00:17.000 00:00:18.000 00:00:19.000 00:00
Clause Transcri	starts picking pears off the tree and he puts the pears into an apron otion non-motion
Gesture #         gestur         gesture 4         gesture 4           [23]         - Gs Hand         R         R         R	gesture 6     gesture 7     gesture 9     gesture 1       R     B     B     B

Screenshot of the audio recognizer panel





The MPIPL is involved in the AVATecH project (Advancing Video and Audio Technology in Humanities Research)

Collaboration of 2 MPI institutes and 2 Fraunhofer institutes

Project website: •http://www.mpi.nl/avatech

Defines a Component Interface Specification, implemented in ELAN's extension mechanism for recognizers

•2 major data types as input for and output from recognizers "TIER" and "TIMESERIES"

•Metadata file format (CMDI) for configuration and parameter specification of recognizer components





Possible use case scenarios:

•Segmentation of speech per speaker by automatic recognition, the user fills in the text

•Segmentation of video into gesture units, the user applies labels

•Import unaligned transcript text, a recognizer aligns the text to the speech signal

•Apply a recognizer to one media file, adjust parameters, save the settings and run the recognizer on multiple other similar files

Caveats:

- •The recognizers are often developed for a particular platform (OS)
- •Unlimited distribution of recognizer software might not be allowed
- •May only be available as a web- or local area service





Working on a local corpus

Export multiple files as

- Tab-delimited text
- Wordlist
- List of annotations
- •A selection of tiers as new EAF

Edit multiple files

•Add, remove and rename tiers and types, change tier attributes

• "Scrub" transcriptions: remove tabs, new line characters, superfluous whitespaces

## Aultiple file processing



		M	ulti File Editor		
Domain					
Load domain					Save changes to domain files
		Tiers	Linguistic Types	1	
(	Add tier			Re	emove tier
Name	Linguistic Type	Annotator		Participant	Children
A_rh_gestype_R1	Gesture	R1		A	
A_bh_gestype_R1	Gesture	R1		A	
A_lh_gestype_R1	Gesture	R1		A	
A_bh_relation_R2	Bimanual Coordinati 🗘				
A_lh_location_R1	Move	R1		A	
A_bh_gestype_R2	Gesture				
A_lh_gestype_R2	Gesture				
A_lh_trajectory_R2	Move				
A_notes_R2	Notes - free values	, R2		, A	
A_rh_trajectory_R1	Move	R1		A	
A_rh_unit_R1	Move	R1		A	
A_lh_trajectory_R1	Move	R1		A	
A_rh_location_R2	Move				
A_rh_unit_R2	Move				
A_lh_location_R2	Move				
A_rh_gestype_R2	Gesture	1			
A_lh_unit_R2	Move	í l			
A_rh_trajectory_R2	Move	i l			
A notes R1	Notes - free values	R1		A	
A_lh_unit_R1	Move	R1		A	
A_bh_relation_R1	Bimanual Coordinati 🛟	R1		A	

#### Multiple file editor

(contributed by Radboud University, Nijmegen)





Working on a local corpus

•Create and export statistics of multiple files

			Sta	tistics				
nnotations	Tiers							
Statistics Variables								
Tier	Number of	Minimal Du.	Maximal Du.	Average D.,	Median Dur.	Total Anno.	Latency	
Gloss1	191	0.38	4.065	1.504021	1.332	287.268	14.74	-
Gloss2	90	0.28	3.15	1.2127	1.107	109.143	31.708	
Кр	308	0.145	9.2	0.956321	0.722	294.547	4.95	
Pius	356	0.25	9.81	1.375829	1.189	489.795	3.1	
K	127	0.28	5.525	1.008803	0.83	128.118	1.38	14
Ricky	23	0.31	2.82	0.939435	0.605	21.607	257.721	
man outsid	103	0.27	2.665	0.872913	0.629	89.91	31.158	
child 1	13	0.34	2.055	1.128692	0.99	14.673	15.95	
loudspeaker	3	3.069	4.308	3.812333	4.06	11.437	8.353	
boy 1	42	0.38	3.74	1.238738	1.205	52.027	32.988	
child 2	5	0.4	2.06	1.08	0.52	5.4	32.988	
boy 2	2	0.67	1.88	1.225	1.225	2.45	135.166	
boy 3	9	0.52	3.74	1.315222	0.93	11.837	178.301	Ŧ

•Create multiple transcription files, for each media file in a folder, optionally on the basis of a template

•Convert multiple files of format X to EAF

•Apply a sequence of operations to multiple files, e.g. create a tier from overlaps of tier A and B, create a tier from the overlaps of tier C and D, merge the two created tiers, rename the merged tier





#### Structured search in multiple files

🖲 🔿 Search eaf files	
Substring Search Single Layer Search Multiple Layer Search	h
Domain: 4 eaf files	Define Domain
Query History: <	
Mode: case insensitive  regular expression	Clear
Minimal Duration         Maximal Duration         Begin After         End Before	)
phasic = 0 annot. repetitive A	ll Tiers
	ll Tiers 🛟
	ll Tiers 🛟
Find #hits : 87 #annotations with a hit : 87 #annotations investigated : 862 Ready >	
#1  phasic   repetitive     #2          #3	
#1 lphasic  lrepetitive     #2          #3          #1       #2          #2          #2          #2          #2          #3	
#1 lphasici repetitivei    #2          #3          #1 lphasici  repetitivei    #2          #3	
#1 lphasic  lrepetitive     #2          #3          #1 lphasic  lrepetitive     #2          #3	
• •	





Possible improvements: Arbitrary grouping of tiers (now grouping based on tier type or participant)

Specify a group of related keywords in a single search action (e.g. inflected forms of a word)

Store and load complex regular expressions

Richer and friendlier presentation of results and more flexible export of search results

Combined search in annotation content (EAF) and metadata (IMDI).





#### Timeseries viewer



Support for a number of file types (.csv, Praat Pitch- and Intensity Tiers, AVATecH xml, proprietary data glove format) Can be extended by new "file readers" Some data (min., max. or avg. of segments) can be extracted and added to a tier as annotation values

#### Possible extensions:

Extract segments based on patterns in a track, e.g. segments where the values are above or below a certain value, segments containing a steep slope





#### **ELAN** - **LEXUS** interaction

**LEXUS** is a lexicon tool that is being developed at the MPIPL It is a web application (desktop version is planned) Website: <u>http://www.lat-mpi.eu/tools/lexus/</u>

First implementation of lookup and retrieval of values from a **LEXUS** lexicon is planned for the next version of **ELAN** 

	Change Typ	e			×	
	Change Ty	/pe				
Current Types						
Type Name Stereotype Us	se Controlled	DC ID	Time-ali	Referen		
phonetic_tran Symbolic Asso		-				
gestures		-	<ul> <li>Image: V</li> </ul>	,	Н	
gesture_phases Time Subdivisi Ge	esture phases	-	✓	,		
gesture_mean Symbolic Asso		-	,	,		
part of speech Symbolic Asso PC	0S	-	,	,		
ecv-test - C\	√ test	-	✓	,	-	
Add Change Delete Impo Select Type	ecv-test				-	
Type Name	ecv-test					
Stereotype	None				-	
Use Controlled Vocabulary	CV test				-	
ISO Data Category				Browse		
Time-alignable	Time-alignable					
References to Graphics Allowed						
Lexus, NGC import 100 entries, ZinLang						
Ct		lose				

0	Lexicon connection
Enter the inf	Lexicon connection ormation of your workspace:
Type:	Lexus
Description:	a web-based lexicon tool from the MPI Nijmegen, the Netherlands
URL:	http://corpus1.mpi.nl/mpi/lexusDojo/services/LexusService
Username:	SignLincProject
Password:	••••••
	< Previous Next > Finish Cancel





#### **ELAN** - **LEXUS** interaction

A tier can be associated with a certain type of elements in lexicon via its linguistic type: •specify the URL of **LEXUS** (or some other lexicon tool) •select a lexicon

•select a "data category"; one of the properties of an entry

C Lexicon connection
Levicon connection
Lexicon connection
Select the lexicon you want to connect to:
NGC import 100 entries
StructureTest
Taalkunde lexicon in NGT
Signetion
<

Lexicon connection					
Lexicon connection Select the element you want to use:					
Name	Description				
BegripNaam	The Desc of different name				
Gebruik	null				
Info	null				
Uitleg	null				
VariantNaam	null				
ZinKort	null				
ZinLang	null				
id	Identification of an element				
linkerhandvorm1	null				
linkerhandvorm2	null				
mediaF	null				
plaats	null				
rechterhandvorml	null				
rechterhandvorm2	null 😼				
< Previous	Next > Finish Cancel				





#### **ELAN** - **LEXUS** interaction

Building a lexicon while annotating
Starting a lexicon from a wordlist extracted from a set of annotation files
Might become the basis for automatic interlinearization (similar to Toolbox's functionality)

Elan - CNGT0001.eaf
<u>F</u> ile <u>E</u> dit <u>A</u> nnotation <u>T</u> ier Type <u>S</u> earch <u>V</u> iew <u>O</u> ptions <u>W</u> indow <u>H</u> elp
Grid Text Subtitles Lexicon Entry Audio Recognizer Video Recognizer Metadata Controls
Annotation: ALSJEBLIEFT Tier: GlosR S2 Constraints: is Get lexicon entry Get lexicon
00:02:00.578 Selection: 00:02:00.578 - 00:02:01.256 678
$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$





Working with Controlled Vocabularies:

•Support for external controlled vocabularies; (un-editable) vocabularies at an internet or local network location

"Lookup" while you type



•Create a Controlled Vocabulary from a selection of data categories (ISOcat - Data Category Registry <u>www.isocat.org</u>)





Assessing inter-rater agreement is important in some types of research

Some users would like to be able to compare the ratings of two annotators in the tool in which they have the annotations

Some preliminary work has been done: the Compare Annotators function

Problem: no consensus on how to compare time aligned annotations (mostly Cohen's kappa is used)

First tier: Second tier Average ag	A_I r: A_I reement: 0.7	rh_trajectory_ rh_trajectory_ '745485	.R2 .R1					
A_rh_traject.	Begin Time	End Time	A_rh_traject	Begin Time	End Time	Overlap	Total extent	Overlap / To
shift	1580	2900	shift	1550	2960	1320	1410	0.9361702
phasic	11730	14270	phasic	11680	14180	2450	2590	0.9459459
stopped/hol	14280	18880	stopped/hol	14190	19100	4600	4910	0.9368635
phasic	27910	31230	phasic	28008	31168	3160	3320	0.9518072
phasic	31230	31720	phasic	31178	31878	490	700	0.7
repetitive	31720	36880	repetitive	31888	37008	4992	5288	0.9440242
repetitive	39320	44640	repetitive	39148	49808	5320	10660	0.4990619
stopped/hol	44640	49720	-	-	-	0	0	0.00
and a state of a	63470	65230	repetitive	63510	65170	1660	1760	0.9431818
repetitive								0.0004007





Co-reference annotation

Allow references from one annotation to multiple other annotations

Has been on the to-do list for a long time

Really necessary for some types of research







ELAN is an annotation tool for time-based media

Some users like to use it for images as well (and found workarounds to achieve that)

A preliminary solution to annotate images more conveniently has been added in the latest version

This might be extended with the option to specify a sequence of images as source for annotation, like pages in a book, where each image has a duration of n ms.

2D annotation might be taken up again...





Currently the program optimizes for precision rather than for speed.

The Segmentation mode could be improved and be better integrated in the application

Special input modes (text entry modes) are required to speed up transcription



Provide better input methods, customizable virtual keyboard (and if possible better rendering of some scripting systems)

Included in an ELAN distribution are a number of virtual keyboards (created by GATE) and some lookup lists by MP

On Windows combined use with Keyman works for some, but not for all

A new platform independent, integrated, customizable virtual keyboard framework would solve some problems

Some operating systems provide better rendering engines than what Java provides; could try to use those

	一些	
	於	
	五	
	物	:00:"
	務	
_	武	
	誤	hove
	嗚	0.0
	舞	
	惡	2000
· · · · ·		_bo:c
2003	4	
eyboard ma	p	

000	Russian (TAWERTT (Phonetic)) Reyboai	ru map	
¬ ю !	$1 \begin{array}{c} *\\ 2 \\ 2 \end{array} \begin{pmatrix} f\\ 3 \\ 3 \\ 3 \\ 4 \\ 4 \\ 5 \\ 5 \\ 5 \\ 6 \\ 6 \\ 6 \\ 7 \\ 7 \\ 8 \\ 8 \\ 9 \\ 9 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0$	- <sup>+</sup> _= ч Е	BackSpace
Tab	$ \begin{array}{c} Q \\ q \end{array} \hspace{0.5cm} \stackrel{W}{\hspace{0.5cm} w^{B}} \hspace{0.5cm} \stackrel{E}{\hspace{0.5cm} e^{B}} \hspace{0.5cm} \stackrel{R}{\hspace{0.5cm} r} \hspace{0.5cm} \stackrel{T}{\hspace{0.5cm} t} \hspace{0.5cm} \stackrel{Y}{\hspace{0.5cm} y^{H}} \hspace{0.5cm} \stackrel{U}{\hspace{0.5cm} u} \hspace{0.5cm} \stackrel{I}{\hspace{0.5cm} u} \hspace{0.5cm} \stackrel{O}{\hspace{0.5cm} o^{O}} \hspace{0.5cm} \stackrel{P}{\hspace{0.5cm} p^{\Pi}} \end{array} $	{ш}	ц∣э
Caps Loc	$k \begin{bmatrix} A \\ a \\ a \end{bmatrix} {}^{S}_{s} c \begin{bmatrix} D \\ d \mathcal{A} \end{bmatrix} {}^{F}_{f} \phi \begin{bmatrix} G \\ gr \end{bmatrix} {}^{H}_{h} x \begin{bmatrix} J \\ i \end{bmatrix} {}^{H}_{k} K \begin{bmatrix} L \\ k \\ \kappa \end{bmatrix} {}^{I}_{j} $	. @	Enter
Shift	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	~ #Ъ	Shift
Ctrl	Alt	Alt	Ctrl







#### •Improve media support, especially on Linux JMF is old software and not maintained since 2002?

•New media solution on Windows to overcome a number of

codec problems

The Four Humoration Tiel Type Zearch Alew		
CAM 2	Filters in the graph.	<u>×</u>
	Filter 1: Video Renderer Pin 1: VMR Input0 Filter 2: Direct Sound Renderer Pin 1: Audio Input pin (rendered) Filter 3: MPEG1 Video Decoder Pin 1: Input Pin 2: Output Filter 4: MPEG1 Audio Decoder Pin 1: XForm In Pin 2: XForm Out Filter 5: MPEG1 Splitter Pin 1: Input Pin 2: Video Pin 3: Audio Filter 6: Source Pin 1: Output	
A =\$0.000 00:00:01.000 00 ↓  A dog was walking a	ici	

Media player for Windows with codec listing

•Connect to (command-line) audio/video editing applications





Many researchers use several tools for different tasks

2 main approaches:

Have specialized tools for specialized tasks and transfer data from one tool to the other

Integrate or copy the functionality of one tool (or a few tools) in the other tool

E.g. many users combine Toolbox (or FieldWorks), Praat and ELAN; some are happy, others would like a "tool merge"





### Thanks

#### http://www.lat-mpi.eu/tools/elan

E. Auer, H. Brugman, G. Gulrajani, A. Klassmann, A. Koenig, M. Kramer,M. Pippel, A. Russel, H. Sloetjes, A. Somasundaram, H. SpenkeM. Blokpoel, M. Hulsbosch, J. Lemein