LANGUAGE DOCUMENTATION & CONSERVATION

Vol. 1, No. 2 (December 2007), pp283-289
http://nflrc.hawaii.edu/ldc/

EUDICO Linguistic Annotator (ELAN)

from Max Planck Institute for Psycholinguistics

Reviewed by ANDREA L. BEREZ, *University of California, Santa Barbara*

1. INTRODUCTION. ELAN was developed by the Max Planck Institute for Psycholinguistics (MPI) in Nijmegen, The Netherlands, and the Java source code is available for noncommercial developers. This review is based on ELAN 3.1.0 for Macintosh; Windows and Linux are also supported. ELAN is freeware.

In what follows, I provide an overview of the stated purpose of ELAN and describe and comment upon its various functions, which include annotation, searching and exporting. The review ends with recommendations for using the program.

2. OVERVIEW. ELAN is designed for annotating digital audio and video. Originally developed for analyzing sign language and gesture, it can be used by "everybody who works with media corpora, i.e. with video and/or audio data, for purposes of annotation, analysis and documentation" (Hellwig and Van Uytvanck 2007:1). In a nutshell, ELAN allows users to link annotations to media streams, to link annotations to other annotations, and to set up user-defined relationships between annotation tiers. The ELAN interface displays up to four video files and/or a waveform (audio file), and allows a number of ways to view the annotations. ELAN supports a range of character sets, plus import from and export to multiple formats (including Shoebox), and has a growing search facility.

My own use of ELAN since 2004 has primarily been to align audio of Dena'ina and Ahtna Athabaskan traditional narratives to text annotations, although I have also used it to annotate video+audio media streams, and experimentally in the study of English discursive interaction.

3. USING ELAN.

3.1. CREATING NEW PROJECTS. The ease of using ELAN is variable: new users may find the initial setup of files difficult, as it involves mastering some potentially confusing concepts as soon as a new file is created. However, once the file is configured properly, the ELAN interface is fairly intuitive and offers a broad range of ways to manipulate both the data and the view of the data. Here I describe the basic procedure for setting up a new ELAN file.

The first step involves importing up to five media files (0-4 video files as .mpg or .mov, and/or one audio file as .wav). I recommend, when working with video, extracting a .wav audio file from your video and importing it into ELAN alongside your video file. ELAN displays the waveform when a .wav is available, making it easier for a user to visualize and "grab" a chunk of audio for annotation. Conveniently, ELAN allows users to synchronize multiple video files that may accidentally be slightly offset. In any case, the first of multiple video files to be imported into an ELAN document is the one that ELAN will link to annotations.

E-ISSN 1934-5275

3.2. IMPORTING FROM OTHER SOFTWARE. ELAN also allows imports of transcriptions created in other programs that can then be linked to the media within ELAN. ELAN accommodates files from Transcriber, CHAT, and Shoebox/Toolbox. I have found that import from Transcriber is fairly simple, but import from Shoebox can be tricky, because the Shoebox/Toolbox field markers need to be defined as ELAN tiers (unless you have a Shoebox .typ file to do this for you). Once imported, the Shoebox/Toolbox annotations can then be linked to the media file's time axis. The ELAN manual contains detailed information on importing and aligning Shoebox/Toolbox projects. Once aligned in ELAN, the file can then be re-exported to Shoebox/Toolbox (see *Exporting* below).

3.3. CONFIGURING LINGUISTIC TYPES, TIERS, AND CONTROLLED VOCABULARIES. Annotation tiers in ELAN are of two kinds: *independent* (or *parent*) tiers, and *referring* (or *child*) tiers. Only the former can be linked to the timeline, and the parent-child linking relationship is such that changes made on a parent tier will also affect its child tiers, but not vice versa.

Before creating tiers you must first configure *linguistic types* for your project. These are stereotypes that dictate the properties of the tiers you will later assign to those types. For example, a parent tier can be divided into smaller time-based segments; thus the tier containing these smaller segments is assigned to the linguistic type *time subdivision*. Other available types include *symbolic subdivision* (a division of the parent tier that is not linked to the time axis, like morphemes within a word) and *symbolic association* (used for glosses or free translations when the parent tier is not further subdivided into meaningful units).

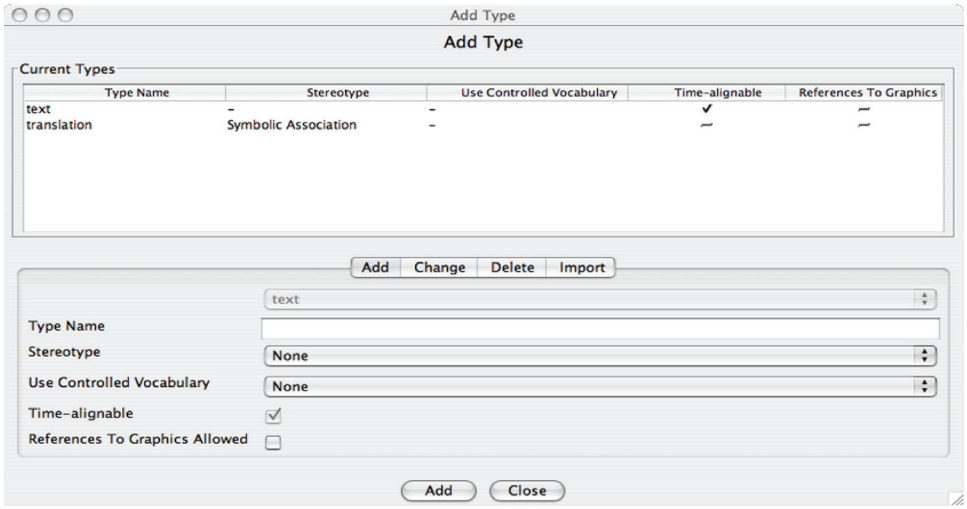

FIGURE 1: "Add linguistic type" dialog box

After creating all needed linguistic types for a given annotation file, the next step is to create the tiers themselves. One of the best features of ELAN is the extreme flexibility users are allowed in creating tiers.[1] An unlimited number of tiers are allowed, provided the linguistic types and the parent-child relationships between tiers are conceived of correctly. For example, a video with four participants may have multiple tiers assigned to each participant: a parent-transcription tier, a word-level tier, a morpheme-level tier, a morpheme gloss, a word-level translation, and a free-translation tier. Information about gesture, eye gaze, etc., can also be placed on parent tiers, as can comments from the analyst. In a video with, say, four participants, a tier configuration of the kind just described for each speaker could lead to a file containing thirty or more tiers. ELAN conveniently allows type/tier configurations to be saved as templates to facilitate the future creation of files with an identical participant structure (e.g. for sequential recordings made during the same field session or later sessions with the same speakers), as well as the importing of individual tier configurations from other ELAN files.

Determining linguistic types and tiers for a given project is potentially off-putting to new users, as it involves (a) understanding the difference between the available types, and (b) being familiar enough with your media to know *a priori* which types and tiers you need. Reading the manual is key to understanding types and tiers, and the description provided there is much more thorough than in older versions of the manual.

ELAN also allows the configuration of *controlled vocabularies,* which are useful if you often use a limited set of annotation values (specific morpheme glosses, for instance). Controlled vocabularies can be created for each new file, or they can be stored in a template and imported for future use.

3.4. ANNOTATION. Annotation begins with the selection of a segment of the time axis, by using the mouse to highlight the desired section of either the waveform or the Timeline Viewer (see *Interface,* below). Once highlighted, the beginning and end points of the selection can be easily adjusted. Automatic "gapless" annotation is also possible. Once a selection is in the right place, it must be saved by double-clicking on the highlighted area, entering annotation text, and then pressing CTRL+ENTER. (This final step is not particularly intuitive and is a source of frustration among new users.) The beginning and end points of saved selections can be adjusted with the mouse, or the entire selection can be moved to a new location in the timeline.

A number of timesaving features are built into ELAN. Quick "on the fly" segmentation, like that found in Transcriber, is a relatively new feature for ELAN, and is a potential time-saver. This feature allows users to listen to a continuous media stream and rapidly mark

1 This flexibility may prove frustrating to advanced users and software developers who want to create or use other applications built upon data extracted from ELAN documents, such as a facility for displaying multiple aligned texts on the web. Such applications often require that all ELAN files have identical tier structures. This may be difficult when you have different people annotating different files, or even when you have multiple files with different tier configurations. It has been said many times: "The best thing about ELAN is the flexible tier structure, and the worst thing about ELAN is the flexible tier structure." This will not be a problem for users who want to simply access the data within the ELAN window.

rough boundaries of, say, intonation units or speaker turns. Users can decide whether an annotation will be defined with one keystroke (for adjacent annotations) or with two (for nonadjacent annotations). Once the basic annotation boundaries are in place, more precise adjustments can be made with the mouse. Other time-saving features include *tokenizing, filtering, and copying* tiers. Tokenizing automatically creates multiple annotation units on a child tier using a delimiter from the parent tier (e.g., creating word-sized annotation units on a child tier based on a multi-word, space-delimited parent tier annotation). Filtering transfers annotations between tiers that are symbolically associated, while copying transfers the complete content of a tier—and, if desired, that tier's children—to another tier or set of tiers.

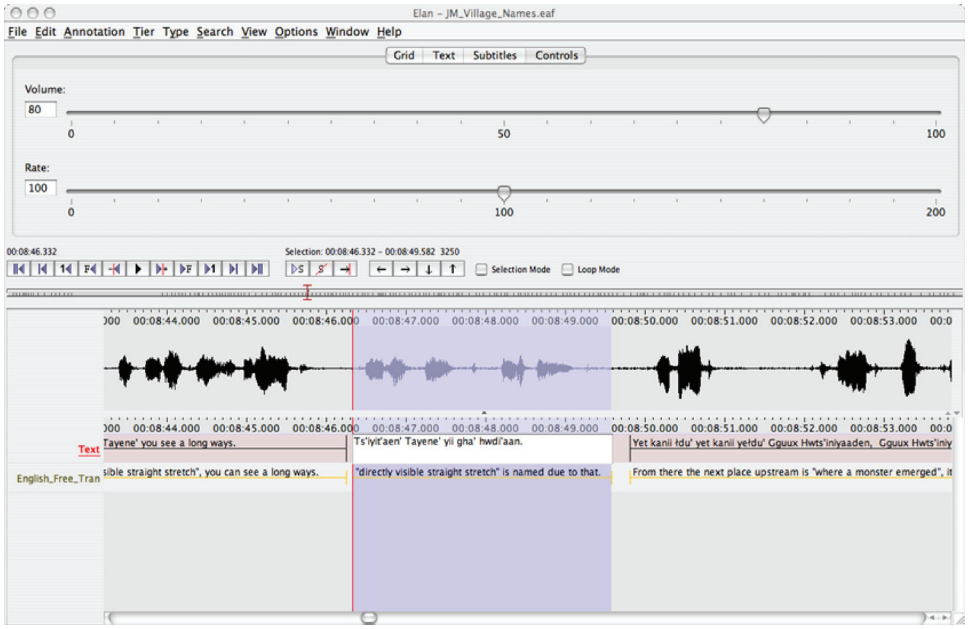

FIGURE 2: ELAN window showing annotation being added

Annotations can be made in a number of different Unicode character sets, including IPA, Chinese (traditional and simplified), Russian, Arabic, and Hebrew. Supported input methods include GATE Unicode Kit, Roman Typographic Root for IPA, or third-party tools like Keyman or Ukelele. Advanced users may want to experiment with adding text directly to the ELAN XML file in a text editor after segmentation of the time axis of the media, as this can be easier than using the ELAN interface. Care must be taken, however, to not disturb the XML structure. Also note that some characters, like ampersand, cannot be read by ELAN and will prevent your file from loading. If this happens, open your ELAN file (.eaf) in a text editor and remove any ampersands.

3.5. INTERFACE. The ELAN window displays the media control buttons, plus a number of different Viewers. Among these are the Video Viewer, showing the video image; the Waveform Viewer, showing the waveform; the Annotation Density Viewer, showing the

entire timeline of the media and the "concentration" of annotations at a particular point along the timeline; the Subtitle Viewer, showing annotations for selected tiers simultaneously with playback; the Grid Viewer, essentially a table showing start and end times of annotations, plus their content; the Text Viewer, showing all annotations of a selected tier as running text; and the Timeline Viewer, showing all tiers and their annotations in a horizontal time-based axis. Within the Timeline Viewer, tiers can be re-ordered for ease of use by click-and-drag without disrupting the dependency relationships between them; unused tiers can also be temporarily hidden. Most Viewers can be switched on and off as desired.

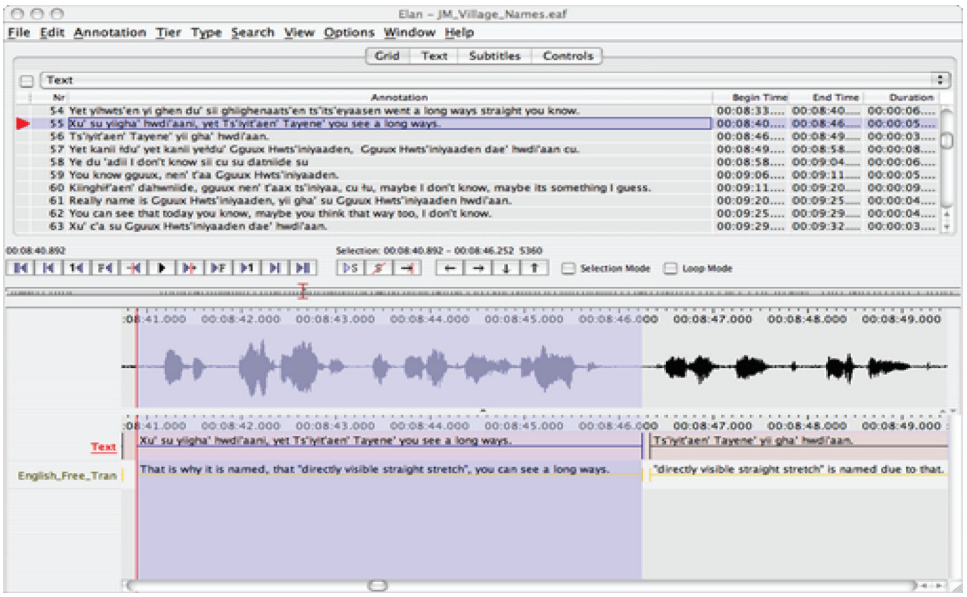

FIGURE 3: ELAN Viewers

4. OTHER FUNCTIONS.

4.1. SEARCH. ELAN's search facility is now a marked improvement over previous versions. It allows searching (and replacing) within a single file with multiple user-designed layered temporal and structural constraints, and accepts regular expression searches. The manual contains a helpful appendix with supported regular expression syntax. The system is flexible enough to allow users to search for tokens as specific as "… all switches from speaker W to speaker K that don't overlap, with gaps of at most 2 seconds" (Hellwig and Van Uytvanck 2007:116). Queries can be saved for future use, and users can jump from the search results to corresponding points in the media. Simple and complex searching over multiple ELAN files is also possible, and the resulting concordance can be exported to a text file.

4.2. EXPORTING. The .eaf file created by ELAN is an XML file, which advanced users may be able to manipulate for other uses with a transformation language like XSLT. But in the likely event that this is not feasible, ELAN allows several options for export. Export into Shoebox is possible, enabling a workflow that potentially involves marking up text

and creating a lexicon in Shoebox/Toolbox, time-aligning text to media in ELAN, and then importing everything back into Shoebox/Toolbox. The re-imported file contains all ELAN tiers and time codes.

Other options for export include tab-delimited text, TIGER-XML (for use with Sympathy, another piece of MPI software for analyzing syntax), CHAT, "traditional" transcripts in .txt format, HTML (with very limited display options and no audio/video), SMIL, QuickTime subtitle text, and even screenshots. For each export type, users can determine which tiers are displayed, whether or not time codes are included, and, to a limited degree, the layout of the exported document. ELAN also offers the possibility of sending selections from the audio file directly to Praat.

5. SUPPORT. Support for ELAN users is substantial. The users' manual, available on the ELAN website, has recently been updated to cover the latest functionality of the program, and explanations are thorough. The website also contains a forum, where a community of users and developers is available to answer questions. I have found this group of people to be helpful and prompt in replying to queries. The website also has a "news" page, with an RSS feed available to keep users informed about software updates.

6. RECOMMENDATIONS. I would recommend ELAN for anyone who wants to align media with annotations and desires flexibility in file structure. However, I offer a few words of caution to new users: the software has a steep learning curve, mainly because of the concepts of types and tiers described above. Also, my own experiences have shown that ELAN is not the best place to work on a "first pass" of a transcript, and in fact, users may want to have a near-final transcript completed before attempting time alignment. This is not because ELAN cannot handle the numerous minute corrections and adjustments that are a natural part of the annotation procedure, but rather because unpracticed users may find the interface too unwieldy. I recommend learning to use ELAN on simple transcriptions of just a tier or two before attempting to align multi-speaker, multi-tier files.

I also want to touch briefly on a misunderstanding about the use of ELAN that I have found to be common among new users during formal and informal training sessions I have conducted: ELAN doesn't "keep your files safe." ELAN creates an archivable document in a long-lasting (XML) format containing information about the contents of a media file. But you still must archive your media alongside an ELAN document with an appropriate repository.

In closing, MPI has vastly improved the functionality, interface, and available support of ELAN since I started using the software in 2004, and is now a versatile, useful piece of software. Anyone who abandoned earlier versions of the software in frustration may want to come back to it now.

Primary function:	Linking complex annotations to time axis of media.
Pros:	Flexible tier structure, archivable XML file, good support and manual, multiple views of data, allows multiple media files, import from and export to numerous common file types.
Cons:	Difficult to learn, interface not visually appealing.
Platforms:	Window, Mac OS, Linux.
Open Source:	Yes. Source code is available from Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.
Proprietary:	No. ELAN is freeware.
Reviewed version:	3.1.0
Application size:	152 KB
Documentation:	User manual, website and forum, all found at http://www.lat-mpi.eu/elan.

References

Hellwig, Birgit, and Dieter Van Uytvanck. 2007. *EUDICO Linguistic Annotator (ELAN) version 3.9.0 manual*. Downloaded 25 July 2007 from http://www.lat-mpi.eu/tools/elan/.

Andrea L. Berez
aberez@umail.ucsb.edu