

The IATH ELAN Text-Sync Tool: A Simple System for Mobilizing ELAN Transcripts On- or Off-Line

Lise M. Dobrin
University of Virginia

Douglas Ross
University of Virginia

In this article we present the IATH ELAN Text-Sync Tool (ETST; see <http://community.village.virginia.edu/etst>), a series of scripts and workflow for playing ELAN files and associated audiovisual media in a web browser either on- or off-line. ELAN has become an indispensable part of documentary linguists' toolkit, but it is less than ideal for mobilizing the transcribed media it allows linguists to create when they have reason to display these materials in non-research settings where linguists are not the primary audience. In conjunction with display of a video or audio file, ETST plays tiers of transcript for overlapping speech, along with optional glosses, and distinguishes speakers with participant codes. Using ETST requires no programming knowledge, but with some such knowledge the tool can be readily customized to suit users' needs. To that extent, ETST is a simple browser-based transcript player that can be used either as is, "out of the box", or as a basis for further development. We hope that ETST will be a helpful addition to documentary linguists' repertoire of digital tools, making it easier for them to share materials with all those who have a stake in their research.

1. Introduction¹ In so many ways, the EUDICO Linguistic Annotator (ELAN) is a documentary linguist's dream tool. Developed by the Max Planck Institute for Psycholinguistics in Nijmegen to facilitate the annotation of digital media (<http://tla.mpi.nl/tools/tla-tools/elan/>), ELAN has revolutionized the practice of corpus creation and analysis in documentary linguistics. ELAN allows users to set up tiers of text and link

¹The authors wish to thank their collaborators at IATH who have helped to make the development of the ELAN Text-Sync Tool possible: Worthy Martin, Daniel Pitti, Amanda Glass, Shayne Brandon, Cindy Girard, Sarah Wells, and Joy Shifflette. We would also like to thank Ed Garrett and David Germano for rewarding discussions about the problem of sharing archival materials with source communities. Finally, we thank an anonymous reviewer for *LD&C* whose generous feedback helped us articulate the ETST workflow more clearly. Funding for the development of ETST was provided by NEH through the NEH-NSF Documenting Endangered Languages partnership's support of the Arapesh Grammar and Digital Language Archive and through a Level I Start-up Grant for a workshop on "Less-Networked Speaker Communities and Digital Language Archives" that was held at the University of Virginia in March 2012. Most importantly, we are deeply grateful to the Arapesh people whose voices can be heard in the AGDLA archive, as well as to Dobrin's many Arapesh friends and relations who encourage and support the work we are doing to record and preserve their traditional language.

them to the time axis of a video or audio recording either directly (as when one provides IPA or orthographic transcriptions for an utterance) or indirectly by association with other tiers (as when one provides additional free translations or morpheme-by-morpheme glosses for those transcriptions). The overlapping talk of multiple speakers that is ubiquitous in human interaction is elegantly handled in ELAN by allowing simultaneous independent linkages between multiple primary transcription tiers and a shared media timeline, in a way similar to the simultaneous and independent linking of multiple tiers for phonological features like tone or nasalization to a central timing tier in autosegmental representations of phonological structure (Goldsmith 1976). The ELAN software is freely accessible and open-source; it runs smoothly on a range of platforms; and the .eaf files it produces are coded underlyingly in XML, the recommended archival format for long-term digital preservation of text (see, e.g., Bird & Simons 2003, E-MELD 2006). The system allows for the transfer of files between ELAN and another piece of software that is widely used for text annotation and linguistic analysis, SIL Fieldworks Language Explorer (FLEX). It also has a robust search mechanism that can run simultaneously across multiple ELAN files. With so many useful features, it is no wonder ELAN has become an indispensable part of many documentary linguists' toolkit.

But ELAN is less than ideal for mobilizing the transcribed media it allows linguists to create when they have reason to display these materials in non-research settings where linguists are not the primary audience. The ELAN interface is optimized for transcription and analysis, not display, so there is no way to rid the interface screen of its many potentially distracting buttons, menu tabs, viewers, and controls. In order to play ELAN files, the program needs to be loaded on a computer and the appropriate steps taken to associate the ELAN file with the right media file. While this might be second nature for a linguist who has spent many hours producing annotations in ELAN, it can be confusing for someone unfamiliar with the program who just wants to hear and view a transcribed recording. ELAN has "a steep learning curve" for new users, and the initial investment of time and focused attention it requires is more than some potential audiences want or are able to make (Berez 2007: 288). As collaboration with source communities becomes more common in linguistics, and as linguists increasingly take seriously the interests of diverse audiences in the materials they produce, the limited accessibility of transcribed recordings in the ELAN format is a problem that begs for a solution. It would help to have tools that make sharing documentary materials with non-ELAN-savvy audiences easier to do.

A number of systems have been developed to make ELAN files viewable outside of ELAN; here we describe two that have the kind of simple, user-friendly interface that makes them attractive for sharing documentary linguistic products with community and non-specialist audiences.² One is the Customizable Presentation of ELAN Doc-

²These are not the only transcript players to be found, just two of the most accessible for non-specialist audiences. Another system that warrants mention is Annex (<http://tla.mpi.nl/tools/tla-tools/annex>), which specifically supports playback of ELAN files archived in the online data repositories (such as DOBES) that are hosted by the Max Planck Institute for Psycholinguistics. The developers of Annex have ported ELAN into a Flash-based interface for interactive display on the web. There is also the transcript player system developed by SHANTI (<http://shanti.virginia.edu/>) for audio and video housed in the Tibetan and Hi-

uments program (CuPED; see <http://sweet.artsrn.ualberta.ca/cdcox/cuped/>). CuPED is a desktop application that allows users to view a media file on a webpage while highlighting the associated transcript in real time. But maintaining CuPED has been a challenge in light of the numerous changes in hardware and software that have taken place since its initial development (COX 2011). For example, CuPED relies on Flash Player, a common third-party plug-in for creating animated web content that, for data-security reasons, is currently being phased out in light of the new capabilities of HTML5, the emerging standard system of protocols used by the web browser industry. As a result, CuPED can run into problems when a user attempts to play the files it creates locally, on the user's computer, without going in and adjusting the security settings in his or her web browser. CuPED also has difficulty playing multiple time-aligned tiers.

Another system that can be used to transform ELAN files and time-linked media into a visually appealing display over the web is the Ethnographic E-Research Online Presentation System for Interlinear Text (EOPAS; see Schroeter & Thieberger 2006).³ EOPAS does not use Flash. However, it is server-based, which means there is no possibility for local playback: the files it creates are not meant to run off-line. Because the processing of ELAN files cannot be disconnected from the EOPAS interface, the web presentations it produces are not easily incorporated into other online environments, like project websites, where one might wish to embed them. EOPAS users submit their materials to the system, which transforms them for online viewing on the EOPAS website at <http://www.eopas.org/> (note that while documentary linguistic materials are stored on this site, it is not meant to serve as a permanent archive). The EOPAS components are all freely available, making it possible to instantiate an EOPAS system entirely on one's own server. However, setting up such a complex system requires a level of expertise in server systems administration that many linguists do not have.

In what follows we present another such system, the IATH ELAN Text-Sync Tool (ETST), which was developed at the University of Virginia's Institute for Advanced Technology in the Humanities (IATH) in order to play the audio and video recordings housed in the Arapesh Grammar and Digital Language Archive (AGDLA; www.ara-pesh.org) in a web browser. One feature of ETST that may be useful to documentary linguists is its functionality either on- or off-line. Off-line display is important for those wishing to share transcribed audiovisual materials with viewers who may have access to power (such as solar) to run a computer but no internet, as is the case for many source communities that reside in remote locations.

ETST visually distinguishes multiple speakers in a transcript by displaying their initials (derived automatically from their ELAN participant labels) at the start of each line, as well as by color-coding those labels and highlighting each ELAN selection as

malayan Library (see, for example, <http://mediabase.shanti.virginia.edu/video/interview-bhikshu-tendzin-p%C3%BCntsok-about-rituals-sangdrok-monastery-2>). However, the SHANTI transcript player is not available for use outside the project's Drupal-based ecosystem. It is also possible to post subtitled video and audio to YouTube; see Cox 2014 for detailed instructions.

³An alternative to EOPAS that will allow users to view ELAN and other file types in the PARADISEC archive is currently under development (Nicholas Thieberger, personal communication, 1/5/2017). The viewer is available on GitHub (<https://github.com/MLR-au/pdsc-collection-viewer/>).

it plays in real time. Tiers of transcript for overlapping speech play simultaneously but are readily distinguished because of these codes. A listing of participants appears at the top of the screen, immediately below which appears the audio or video with an embedded play/pause bar. A viewer can play/pause individual lines by clicking on their line-initial labels. ETST optionally presents a gloss at the bottom of the screen for each line of the transcript as the latter scrolls vertically in the main player window (see Figure 1).

When Timothy Fell in The Latrine

LD · Lise Dobrin SS · Scola Sonin

The screenshot shows a web-based transcript player. At the top, there is a progress bar with a play/pause button and a volume icon. Below the progress bar is a list of transcript lines, each starting with a participant label (LD or SS) and a line of text. The line "Waka yawic yecirokuhen juwebər, yati wa yaren yabəh worib yabəh yecirokuhen." is highlighted. Below the transcript list is a section labeled "Gloss:" followed by a line of gloss text: "SS Nau mi kisim em igo insait nau wasim em long hotwara nau mi lukim nogat nau mi kisim em igo daun long wara nau mi wasim em gen."

Figure 1. ETST display with Cemaun Arapesh transcript and Tok Pisin glosses

Using ETST requires no programming knowledge, but with some programming knowledge the tool can be readily customized to suit users' particular needs, for example by increasing the number of participants it differentiates, providing for multiple gloss tiers per transcription, embedding the presentation frame into another webpage, etc. To that extent, what we present here is a simple browser-based transcript player that can be used either as is, "out of the box", or as a basis for further development. Coded entirely in standards-compliant HTML without reliance on third-party plugins, ETST is maximally efficient, reliable, and compatible with other systems. ETST was also designed to be responsive to different digital environments, working across devices and display resolutions. This makes ETST presentation files viewable on a smartphone or tablet as well as on a computer.

2. How ETST works There are three primary files that make up the ETST system:

1. a Hypertext Preprocessing file (**txt_sync.php**) that takes the XML-based information in an ELAN file and transforms it into HTML so that it can be read by a web browser,
2. a JavaScript file (**txt_sync.js**) that drives the display's interactive features such as scrolling and clickable play/pause, and
3. a Cascading Style Sheet file (**txt_sync.css**) that provides the formatting instructions for the display.

Together these three files, which can be downloaded freely from the IATH ELAN Text-Sync Tool website (<http://community.village.virginia.edu/etst/>), comprise a mere 11 KB of data. When combined with an ELAN transcript and synchronized media in a standard file format (.mp3 for audio or .mp4 for video),⁴ the system produces a derived HTML file that provides the commands for integrating the materials and playing them in a web browser. For materials that are stored on a server and are intended to be made accessible over the internet, the derived HTML file can be generated dynamically, with the PHP script assembling the components at the time of a user's request, i.e., when the user clicks on a link. Alternatively, for materials that are intended to be played off-line, the HTML file can be pre-generated according to the process described below while the user has access to a server, and then saved along with the primary ETST files on a laptop, DVD, or external drive.

3. How to set up an ELAN file for ETST display In order for the ETST system to properly reference an ELAN file, the ELAN file needs to be configured to be readable by the system. For users' convenience, an ELAN template preconfigured for two speakers is included with the downloadable files on the ETST website. While we understand it is unlikely that a preexisting ELAN file created for other purposes will already be configured in exactly the right way, it will often be possible to resave and modify the file so that its tier structure conforms to the system's specifications, because what ETST requires is relatively simple.

ETST makes reference to two different categories of tier, or "linguistic types" in the technical vocabulary of ELAN.⁵ An ETST user's first task is to make sure the relevant ELAN file makes use of those two linguistic types. For primary time-aligned text, the system makes reference to tiers equivalent to ELAN's "default" type, with no pre-defined stereotype; this linguistic type should be given the type name "Transcription". For glosses, the system makes reference to tiers that are dependent on the Transcription tiers in the stereotypic relation of "Symbolic Association"; this linguistic type should be given the type name "Gloss". (See Figure 2.) Note that while ETST requires that each ELAN file have at least one Transcription tier (and up to

⁴Archival materials can be converted into these standard compressed file formats using programs like Audacity or iTunes; there are also online converters available for this purpose.

⁵For an introduction to the linguistic types and tiers of ELAN, see the ELAN manual at <http://www.mpi.nl/corpus/html/elan/index.html>.

eight can be handled by the labeling system as it currently stands), associated Gloss tiers are optional. One advantage of ETST is that tiers formatted according to any other linguistic type will simply be ignored, so it is not necessary to delete extra tiers to get the system to work.

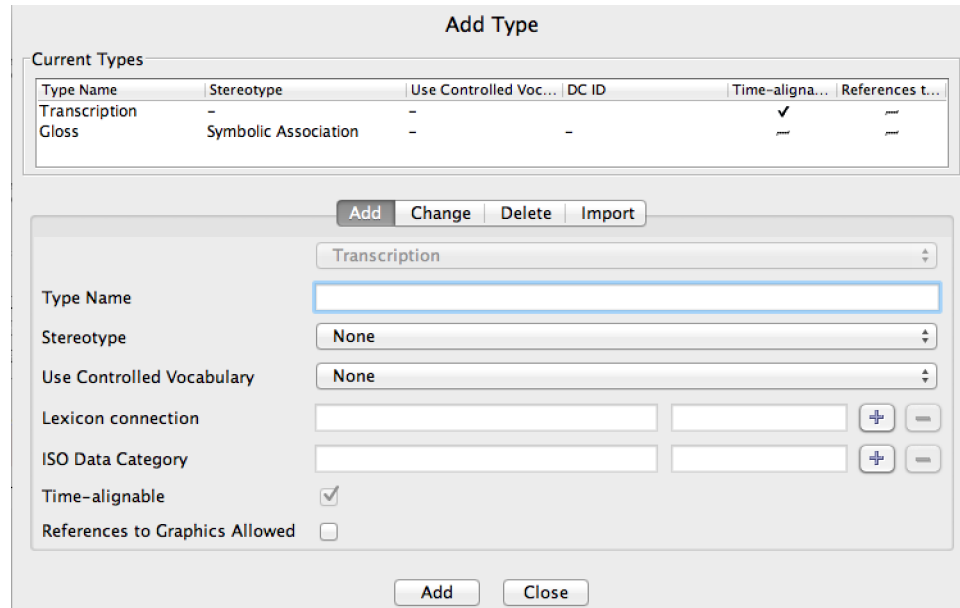


Figure 2. Required type configurations in ELAN

The labeling of ELAN file tiers for use with ETST is similarly straightforward. Each tier, whether Transcription or Gloss, should be given a participant label, and this participant label must be the same for each associated pair of Transcription and dependent Gloss tiers. In general, it will make sense to use the participants' names or titles for the participant labels, since those are what ETST abbreviates to produce the color-coded labels distinguishing different speakers in the display. (If a tier has no participant defined, its contents will still be displayed, but it will have no label.) The tiers themselves can be named however the user wishes. Tier and type names are not case-sensitive (see Figure 3.)

Finally, in order for the generated HTML to function correctly, the ELAN file must be associated with a media file that is (a) in the required standard format (.mp3 or .mp4 – not, e.g., .wav or .avi), and (b) stored in the correct location in the system files (described below).

4. Getting started with ETST The steps outlined below are intended to serve as a “quick start” guide for new ETST users. They allow a user to generate an HTML file that will play locally, without mediation through a server. It is also possible to set up the ELAN Text Sync Tool on a server for dynamic file processing. In order to do this, it is necessary to upload the ETST materials to an internet-accessible directory on a

Add Tier

Tier Name	Parent Tier	Linguistic Type	Participant	Annotator	Input Method	Content Lang...
Lise Dobrin	-	Transcription	Lise Dobrin		English (Unite... -	
Scola Sonin	-	Transcription	Scola Sonin		English (Unite... -	
Scola Sonin TP Gloss	Scola Sonin	Gloss	Scola Sonin		English (Unite... -	

Add Change Delete Import

Lise Dobrin

Tier Name

Participant

Annotator

Parent Tier

Linguistic Type

Input Method

Content Language

Figure 3. Example ELAN tier structure

web server capable of running PHP. For those familiar with HTML and programming for the web, a detailed description of the ETST component files and the principles governing the system is included in the documentation on the ETST website.

In addition to the three primary system files mentioned above, two example files are included on the ETST website: one ELAN file (**example_elan.eaf**) that conforms to the specifications above, and one properly formatted media file (**example_audio.mp3**). A pre-generated HTML file (**txt_sync_example.htm**) is also included as an illustration of what the system produces. Once users are comfortable with the way ETST works and the formats it requires, they can begin to systematically populate the **elan_files** and **media_files** subfolders with their own materials.

1. Go to the ETST website at <http://community.village.virginia.edu/etst/>.
2. Download and unzip the “ETST and Example Files”.
3. Move the ELAN file you wish to display into the **elan_files** subfolder; move its associated media file into the **media_files** subfolder. If you do not have an ELAN file to hand, you can use **example_elan.eaf** in the downloaded example files.
4. Confirm that the media file is in an acceptable format (.mp3 or .mp4) and is correctly referred to by your ELAN file (you can check this by clicking on Linked Files in the ELAN Edit menu).

5. Click on “Upload ELAN for Transformation”.
6. Browse to your ELAN file.
7. Add a Title, then click the “Transform” button to generate an HTML file.
8. Save the HTML file to the `elan_text_sync_tool` folder.
9. Open the HTML file in a web browser to view the results.

The generated HTML file will be given the same name as the uploaded ELAN file. HTML files must be kept directly in the `elan_text_sync_tool` folder (not in either of the subfolders), and associated media files must be kept in the `media_files` sub-folder. For those working locally, the original ELAN files may be kept in the `elan_files` sub-folder or discarded once the HTML file has been generated.

5. Concluding remarks The IATH ELAN Text-Sync System will surely not meet every possible text display need felt by every ELAN user. In its current form it does not accommodate more than one gloss tier per transcription or offer choices in where the gloss displays, nor does it allow for the display of interlinear text or have a built-in mechanism for access control.⁶ But by making the core of an elegant modular display system available, by describing it in a way intended to facilitate its adoption by those without programming knowledge, and by providing detailed documentation for those do have such knowledge and wish to develop the tool further, we hope that ETST will be a helpful addition to documentary linguists’ repertoire of digital tools, making it easier for them to share materials with all those who have a stake in their research.

References

- Berez, Andrea. 2007. A review of EUDICO Linguistic Annotator (ELAN) from Max Planck Institute for Psycholinguistics. *Language Documentation & Conservation* 1(2). 283–289. <http://hdl.handle.net/10125/1718>.
- Bird, Steven & Simons, Gary. 2003. Seven dimensions of portability for language documentation and description. *Language* 79(3). 557–582. doi:10.1353/lan.2003.0149.
- Cox, Christopher. 2011. The ecology of documentary linguistic software development. (Paper presented at the 2nd International Conference on Language Documentation and Conservation (ICLDC), Honolulu, February 11–13, 2011). <http://hdl.handle.net/10125/5239>.
- Cox, Christopher. 2014. Advanced ELAN methods. (Handout distributed in Advanced ELAN Methods course, Institute on Collaborative Language Research (CoLang), Arlington, TX, June 16, 2014).

⁶Although it was not designed to work this way, the system can be constructively “hacked” in order to get glosses, interlinear annotations, notes, or other information to display alongside the transcript in the main viewer window by setting them up as Transcription tiers and giving them identical time selections as the transcribed text.

- Goldsmith, John. 1976. An overview of autosegmental phonology. *Linguistic Analysis* 2(1). 23–68.
- E-MELD. 2006. E-MELD school of best practice: Digitizing text. <http://emeld.org/school/classroom/text/index.html>.
- Schroeter, Ronald & Thieberger, Nicholas. 2006. EOPAS, the EthnoER online representation of interlinear text. In Barwick, Linda & Thieberger, Nicholas (eds.), *Sustainable data from digital fieldwork*, 99–124. Sydney: Sydney University Press.

Lise M. Dobrin
ld4n@virginia.edu

Douglas Ross
dr3f@virginia.edu