To: MesoSpace team
Subject: ELAN - a test drive – version 3
From: Jürgen (v1), with additions by Ashlee Shinn (v2-v3)
Date: 9/19/2009

This document - I witnessed (from a decidedly peripheral position) the development of ELAN during my final years at the Max Planck Institute, but didn't try to install and use it until 2004. Suffice it to say, we weren't ready for one another then. This week, I tried again. I downloaded and installed ELAN 3.5.0 for Windows, released two short months ago. I did so searching for a tool to transcribe the MesoSpace referential communication sessions - the sessions recorded with the *Chunches* and Ball & Chair. Until now, whenever I videotaped sessions that I expected to transcribe, I made a simultaneous audio recording and based the transcription on the latter (using Transcriber since 2004). This year for some reason I decided to skip the audio recordings and go with video only. If you do this, there are programs that you can use to convert the digital video file into a digital audio file (see below) and then use again Transcriber or your tool/method of choice. However, I decided to give ELAN another try,[1] and the results are passable enough for me to be ready to endorse ELAN now for others looking for a transcription tool. So if you are looking for such a tool, this document is for you, but if you are not interested in using transcription/annotation software, please skip. Given my familiarity with Transcriber and because, and partly despite, the documentation that comes with the installation package, I was able to figure out the basics of using ELAN in about two to three hours. In the process I started taking notes - using software involves sufficient amounts of declarative knowledge to make writing stuff down helpful in internalizing it. Then I got carried away a little. The result is the present document.

History, functionality, distribution - ELAN is a program for annotating, transcribing, and coding digital audio- and video recordings. It has been designed and implemented continuously for the last eight years or so by a team of the Technical Group at the Max Planck Institute for Psycholinguistics. ELAN is the successor of the MediaTagger video annotation software for the Mac developed in the mid 1990s for the MPI Gesture Project. The goals behind the transition from MediaTagger to ELAN were the following: (a) Make the new product work on a larger number of platforms. There are currently releases available for download from the MPI website for Mac, Windows, and Linux. (b) Make this thing useful

---

[1] One aspect of the Yucatec *Chunches* sessions that makes me prefer a video annotation tool is the rampant ambiguity involved in Yucatec meronym assignment. Coding the meronym assignment becomes a lot easier when you can tell which part of a *Chunche* the speaker is referring to by how they are holding the *Chunche* and what they are looking at while talking. In addition, gesture cues can be important for figuring out what frame of reference speakers are using. Of course, all of this information is available (provided you video-taped the sessions) whether or not you use a video annotation software; it's just that the software makes checking so much easier and faster.

not just for gesture research, but also for field workers and psycholinguists primarily interested in speech and anybody who needs to annotate, transcribe, or code audio or video recordings in linguistics and neighboring disciplines. Software for annotating digital audio files has been around for a while and is probably better known than MediaTagger outside the gesture research world and the Max Planck universe. Probably the program with the widest distribution at present is Transcriber. From this perspective, the goal of the ELAN team was to come up with something that would offer the same capabilities as Transcriber, but would not have some of the limitations of that software (e.g., Transcriber does not allow the use of extended-ASCII or non-ASCII character sets), and would be usable with both audio and video files (Transcriber works only on audio files).

The idea behind programs like MediaTagger, Transcriber, and ELAN is a construct that combines one or more media files with one or more text files, spreadsheets, or databases, linked via a time code, so that one can assign strings of text (etc.) - the annotation, transcript, or coding - to particular segments of the media file(s). Play back the segment and you can view and edit the annotation on screen - the usefulness of that should be obvious to anybody who has ever tried to transcribe a recording. Executable installation packages for the various platforms are freely downloadable from the website of the MPI (go to www.mpi.nl and look under "Tools"). The program comes with an extensive manual accessible via the "Help" menu. It is not always easy - for me, anyway - to find in this manual what you are looking for or to understand what you are being told. There are a number of reasons for this. ELAN is substantially more complex and has substantially more capabilities than its competitors. And it was designed and programmed by a large team whose credited members do not include any linguists (although linguists were in fact consulted extensively during the design phase); so the manual is written from the perspective of software developers and computer scientists and not from that of users. Finally, ELAN is non-commercial software. It's a work in progress; not everything works yet or works the way it was meant to.

Displaying the waveform - ELAN only does this for .wav files. Once you have the .wav file, you can open the .mpg and .wav files simultaneously and so use the waveform as a convenient frame of reference for segmenting the speech stream the way it's done in Transcriber (and Praat, etc.). But you have to *have* a .wav file first. There are programs that convert .mpg to .wav, but ELAN itself doesn't do this.

The term "annotation" - In order to make sense of the menus and help text/manual of ELAN, it is important to understand a terminological convention that isn't necessarily intuitive (in my view) and is apparently nowhere documented. Each **tier** of your ELAN transcript (if the concept isn't familiar from Transcriber (etc.) already, you can think of the tiers as separate transcripts linked to the same media file(s), e.g., one per speaker/participant; cf. below) involves a segmentation of the video and/or audio signal of the recording into chunks. Each of these chunks has a beginning and end point (which are usually,

though not necessarily, "time-aligned", i.e., defined with respect to particular time codes of the media file(s) under transcription) and may contain bits of transcription that represent some of the information in the recording during the interval the chunk corresponds to. The ELAN people call each such chunk or segment of a tier an **annotation** - whether or not it actually contains any transcription. So just remember that by *annotation*, they mean any segment of a tier defined in terms of a beginning and end point (see below on how to do this).

Creating tiers - Of course you can use tiers for all sorts of stuff. I only transcribe speech data (no gesture etc.) and don't do any phonetic, morphological, or what-have-you coding/annotation/analysis as part of the transcription itself; so for me, it is useful to define one tier per participant in a session and that's the extent of it. To create a tier, go to the Tier menu, select "Add New Tier", give the tier a name (which can be the same as the identifier you use for the participant), enter the participant whose speech is going to be transcribed on this tier, identify the annotator (you?). You also get to specify a "Linguistic Type". Of course what the ELAN folks mean by *Linguistic Type* is no such thing, but is in fact a type of tier. This is useful for example if you want to transcribe both speech and gesture, or have separate tiers for phonetic and morpho-phonological transcription, etc. You can define your own "Linguistic Types" in the "Type" menu. However this doesn't seem to work yet (I can define types at will but they don't show up for selection when adding new tiers). In any case, as long as you only transcribe speech and only use one type of transcription (say, phonological), you don't need this. Finally, you get to select a "Default Language" for the tier. This is important because it determines the character sets and character input options you have at your disposal for the new tier. See below.

Aligning time codes throughout tiers- If you want to base the segmentation of all your tiers on one primary tier, or just make the time codes of one tier dependent on another, you must create a 'parent/child' relationship between the tiers. But before you can do that, you need to make it so that the stereotype of the linguistic type is not time-aligned independently (in Elan terms "time-alignable"); you will need to create a new linguistic type ("add type" from 'Type' menu) with the stereotype "symbolic association". All you have to do is give it a name and chose the stereotype. There must be a tier which *is* time-alignable, i.e. the one that the other tiers will be dependent on for their segments. If you want to align the segments of an existing tier, you must change the linguistic type of the tier and then make it the 'child' of another tier. Oddly, this will create a separate tier, identical except for the fact that it will adopt its parent's segmentations. Otherwise, to do this for a new tier, define the linguistic type with the new type you have created. And here's where the not-so-obvious workings of Elan show their full glory: you must define the parent tier of the new tier before the linguistic type you have created is displayed and can be selected. Now you have a new tier, but it will not have time segments until you go to add an annotation; each annotation will automatically copy the segmentation of its parent.

It should be noted that a tier's default setting is parent-less, so in order to give it one, you select "change parent of tier" in the tier menu, and choose which tier you want to be its parent. During this process you are also able to define/create a new linguistic type; choose "symbolic association" and give it a name; beyond this, nothing need be done. The result of this sort of a change is an identical tier that now adopts its segmentation from its parent. When exported as a tab-delimited text (with the option of separate column for each tier), your file will then be able to have a single row following the value of whatever your parent tier is.

Segmentation - The segmentation of tiers into annotatable chunks synchronized to the media file(s) under transcription is referred to as **Creating New Annotations** in ELAN, in line with the weird terminological convention mentioned above. As in Transcriber, it is possible to define initial segmentations just by playing the media file(s) and hitting a key whenever you want to insert a breakpoint. Except none of this is called that way. Go to "Tier" > "Annotation"; select the tier you want to segment; select the segmentation method - "One keystroke" if you want every keystroke to define a breakpoint as in Transcriber, "Two" for defining the beginning and end of non-adjacent "annotations", or one breakpoint at every fixed time interval; click the play button *on the "Segmentation Preview" controls at the bottom of the "Segmentation" menu*!; hit a key every time you want to insert a new breakpoint, starting from the current cursor position; *and then when you're done defining breakpoints click "Apply" to actually see the breakpoints you just defined*! On my laptop, this routine runs kind of choppily. So you get lots of breakpoints that aren't really where you want them, and you need to do some adjusting, which is a lot less user-friendly in ELAN compared to Transcriber (see below). The other method of defining a segment (or an "annotation") is by **selecting an interval** (click on the time line or the tier you want to segment at the position of the left boundary, hold the left mouth button, and drag the cursor to the intended right boundary) and choosing "Annotation" > "New Annotation Here" (or right-click on the selected segment and select "New Annotation Here" from the pop-up menu). This presupposes selecting one tier as the **active tier**, for example by double clicking on the name of the tier - the "New Annotation" segment will be defined only for the active tier. You can **play the selected segment** of the recording by clicking, not the Play button underneath the video monitor, but the "↳S" button on the controls to the right of the main media controls, underneath the displays toggled by the "Grid"/"Text"/"Subtitles"/"Control" taps. To **play an "annotation"**, select it by clicking into its representation in the **timeline viewer** (the lower part of the ELAN display presents a separate timeline for each tier) and use the "↳S" button. To **extend or reduce a selection**, hold the Shift key down and click on the time line where you want the changed breakpoint to end up. The results of this can be surprisingly difficult to predict ;-) If you use the "New Annotation Here" function, this automatically calls up the so-called **Inline Edit box** used to enter the text of "annotations" (see below). Just click outside it to make it go away if you're not yet ready to enter text. To **change an "annotation"** segment (the equivalent of moving a breakpoint in Transcriber simply by dragging it with the mouse while holding down CNTR; unfortunately, ELAN does not have any quite

as intuitive way of doing this), you have two options. You can use either the "New Annotation Here" function on an interval that overlaps with the segment you want to change - this will extend or split the old segment; or you select the "annotation" you want to change by clicking into its representation in the timeline viewer, pressing and holding down the Shift key, extending or reducing the selected segment in the desired direction by left-clicking on the timeline where you want to changed breakpoint to be, right-clicking, and selecting "Change Annotation Time" from the pop-up menu. In Transcriber, you can **merge segments** simply by deleting the breakpoint that divides them; in ELAN, you have to select a segment that corresponds to both former segments and use the "New Annotation Here" function on the selection.

Annotating a segment - As in Transcriber, a new file is created with the information about the segmentation you define. The program will ask you where you want to save this new file. When you assign text to particular segments - transcription, annotation, or coding - it will be stored in that file. You can later export the transcript etc. in a format that more or less suits your needs; see below. If you double-click on a segment ("annotation"), the **In-line Edit box** appears; this allows you to enter text. Crucially, **the text won't actually be assigned to the "annotation" until you hit CNTRL-ENTER!** Once "committed", the text is displayed in two places, above the interval icon that represents the segment in the timeline viewer and also in the "Subtitles" viewer to the left of the video monitor. Click on the "Subtitles" tap and select the tiers you want to be displayed.

There are several options for entering special characters. These are documented in the manual. Go to "Help" > "Help Contents" > "Users Guide" > "Annotations" > "How to enter annotations" > "Entering annotations in different character sets". Unfortunately, if you're me, you won't be able to figure out how any of these options are going to help you enter something as pathetic as vowels with acute and grave accents to represent the tones of Yucatec. There is text and diagrams about "diacritics" that may be relevant; but I just can't work out what any of that is supposed to mean ;-) However, it turns out that you can use the ALT-plus-numeric-key-pad-combination of MS DOS days long past. You can do that in Transcriber as well, but it seems to wreak havoc there.

Export transcription - There are oodles of export options ("File" > "Export As"). To generate something that looks like a Shoebox text output, with each "annotation" (see above) on a separate line starting with a code that identifies the tier, use the "Traditional Transcript Text" option. Be sure to select all tiers you want to be included in the export and check the "Include Tier Labels" box.

Jesus saves! - Everybody knows the old joke about Jesus and Lucifer competing on who is better at computer stuff. Don't forget the punch line! Earlier versions of ELAN were reportedly pretty unstable. This new one hasn't crashed on me yet; but as of the writing of this doc, I haven't started serious transcription yet. So *caveat emptor*!

Comment [t1]: Continue here.